

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 14, 2003, 14:14:42 ; Search time 230 Seconds  
(without alignments)

2084.096 Million cell updates/sec

Title: US-09-374-967-1

Perfect score: 1086

Sequence: 1 atgaagcccttcattc-tgt.....agcctgagatcgtcatgtga 1086

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapex: 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA.\*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/2/ina/FCRUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	399.6	36.8	1559	4	US-09-786-240-17
2	74.2	6.8	1535	3	US-09-032-372-8
3	73.2	6.7	787	4	US-09-007-119-20
4	58.4	5.4	7902	4	US-08-961-527-112
5	55.2	5.1	690	3	US-09-024-023-1
6	55.2	5.1	690	4	US-09-531-111-1
7	53.6	4.9	287	3	US-09-024-023-3
8	53.6	4.9	287	4	US-09-531-111-3
9	52.8	4.9	2634	1	US-08-196-218-31
10	52.8	4.9	2634	1	US-08-681-953-31
11	46.2	4.3	678	4	US-09-370-838-155
12	42.2	3.9	891	4	US-09-107-532A-2711
13	41.8	3.8	702	4	US-09-328-352-3686
14	41.6	3.8	6854	4	US-09-194-905-7
15	40.4	3.7	3571	4	US-09-221-017B-904
16	40.2	3.7	28804	2	US-08-592-874-1
17	40.2	3.7	28804	3	US-09-096-942-2
18	40.2	3.7	28804	3	US-09-096-867-2
19	39	3.6	960	4	US-09-252-991A-6944
20	39	3.6	1272	3	US-09-036-987A-28
21	39	3.6	1272	3	US-09-370-700-28
22	39	3.6	1272	4	US-09-603-207-28
23	39	3.6	1443	4	US-09-252-991A-6752
24	39	3.6	1695	4	US-09-252-991A-6801
25	38.2	3.5	164976	4	US-08-916-421B-1
26	37.4	3.4	5144	2	US-08-487-890A-105
27	37.4	3.4	5144	2	US-08-478-435-105

#### ALIGNMENTS

##### RESULT 1

US-09-786-240-17  
; Sequence 17, Application US/09786240  
; Patent No. 6558935  
; GENERAL INFORMATION:  
; APPLICANT: INCYTE PHARMACEUTICALS, INC.  
; APPLICANT: TANG, Y. Tom  
; APPLICANT: CORLEY, Neil C.  
; APPLICANT: GUEGLER, Karl J.  
; APPLICANT: BAUGHN, Mariah R.  
; APPLICANT: LAL, Preeti  
; APPLICANT: YUE, Henry  
; APPLICANT: HILLMAN, Jennifer L.  
; APPLICANT: AZIMZAI, Yalda  
; TITLE OF INVENTION: HUMAN TRANSFERASE PROTEINS  
; FILE REFERENCE: PF-0592 PCT  
; CURRENT APPLICATION NUMBER: US/09/786,240  
; CURRENT FILING DATE: 2002-03-12  
; PRIOR APPLICATION NUMBER: 09/150,657; unassigned; 09/186,779; unassigned; 60/133,642  
; PRIOR FILING DATE: 1998-09-10; 1998-09-10; 1998-11-04; 1998-11-04; 1999-05-11  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: PERL Program  
; SEQ ID NO 17  
; LENGTH: 1559  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. 6558935 2682663CB1  
US-09-786-240-17

Query Match 36.8%; Score 399.6; DB 4; Length 1559;  
Best Local Similarity 61.1%; Pred. No. 6.8e-121;  
Matches 664; Conservative 0; Mismatches 419; Indels 3; Gaps 1;

Qy	1	ATGAAGCCCTTCATCTTCTCGGGGGTTTCGAAACCCGCTTTCGGGCTTGTACTGTGAGC	60
Db	213	ATGAAGCCACTGATCTTAGTGGGGGCTATGGACGGCGCTACGGCGGTGACGCTGAGC	272
Qy	61	TTCCCGAAACCCCTCGTGGATTTCGAAACAGCCCATGATTCTGCACACATCCAACT	120
Db	273	ACCCGAGCCACTGGTGGACTTCTGCAATAGCCCTTCTGCTGCACCAAGTGGAGCGG	332
Qy	121	TTGAAGAAGTTGGGGTCACAGAGTGGTTTGGCTATCAACTATCGCCAGAGGTAATG	180
Db	333	CTAGCCGGCGGAGGGTGGACCACTGATCTCGCCGTCGAGTACATCTCGCAGGTGTG	392
Qy	181	ATTATTTCTTGAAGGACTTTGAGGATTAAGTTGGCATCACAATTACATGTCCCAAG	240
Db	393	GAGAAGGAATGAAGGCACAGGAGGCTGGGAATCCGAATCCGAATCTCCATGTCCCAT	452



Db 533 TGCATCGTTGAGATCCACACACACGAGGTATTGCATCTGTGGAGAACCCAGCACA 592  
Qy 499 TTTGTGGTAACAAGATCAATGCTGGGATTACTT 533  
Db 593 TTTATCATGTGATCATCAATCAACTGGCGCATCTACCT 627

RESULT 3  
US-09-007-119-20/c  
; Sequence 20, Application US/09007119C  
; Patent No. 6300541  
; GENERAL INFORMATION:  
; APPLICANT: Lightfoot, David A.  
; APPLICANT: Gibson, Paul T.  
; APPLICANT: Merkem, Khalid  
; TITLE OF INVENTION: Soybean Sudden Death Syndrome Resistant Soybeans,  
; TITLE OF INVENTION: Soybean Cyst Nematode Resistant Soybeans and Methods of  
; TITLE OF INVENTION: Breeding and Identifying Resistant Plants  
; FILE REFERENCE: Sou Illinois 1268/2 Sequence Listing  
; Patent No. 6300541  
; CURRENT APPLICATION NUMBER: US/09/007,119C  
; CURRENT FILING DATE: 1998-01-14  
; EARLIER APPLICATION NUMBER: 60/035,335  
; EARLIER FILING DATE: 1997-01-14  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 20  
; LENGTH: 787  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)  
; OTHER INFORMATION: (a or c or g or t/u)  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (3)  
; OTHER INFORMATION: (a or c or g or t/u)  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (11)  
; OTHER INFORMATION: (a or c or g or t/u)  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (15)  
; OTHER INFORMATION: (a or c or g or t/u)  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (129)  
; OTHER INFORMATION: (a or c or g or t/u)  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (152)  
; OTHER INFORMATION: (a or c or g or t/u)  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (162)  
; OTHER INFORMATION: (a or c or g or t/u)  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (173)  
; OTHER INFORMATION: (a or c or g or t/u)  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (181)  
; OTHER INFORMATION: (a or c or g or t/u)  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (228)  
; OTHER INFORMATION: (a or c or g or t/u)  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (526)

; LOCATION: (265)  
; OTHER INFORMATION: (a or c or g or t/u)  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (292)  
; OTHER INFORMATION: (a or c or g or t/u)  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (329)  
; OTHER INFORMATION: (a or c or g or t/u)  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (335)  
; OTHER INFORMATION: (a or c or g or t/u)  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (339)  
; OTHER INFORMATION: (a or c or g or t/u)  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (351)  
; OTHER INFORMATION: (a or c or g or t/u)  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (363)  
; OTHER INFORMATION: (a or c or g or t/u)  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (384)  
; OTHER INFORMATION: (a or c or g or t/u)  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (387)  
; OTHER INFORMATION: (a or c or g or t/u)  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (396)  
; OTHER INFORMATION: (a or c or g or t/u)  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (400)..(401)  
; OTHER INFORMATION: (a or c or g or t/u)  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (421)  
; OTHER INFORMATION: (a or c or g or t/u)  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (428)  
; OTHER INFORMATION: (a or c or g or t/u)  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (432)  
; OTHER INFORMATION: (a or c or g or t/u)  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (462)  
; OTHER INFORMATION: (a or c or g or t/u)  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (475)  
; OTHER INFORMATION: (a or c or g or t/u)  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (514)  
; OTHER INFORMATION: (a or c or g or t/u)  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (521)  
; OTHER INFORMATION: (a or c or g or t/u)  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (526)

Query Match	6.7%	Score 73.2;	DB 4;	Length 787;
Best Local Similarity	56.0%;	Pred. No. 8.1e-14;		
Matches 227;	Conservative	0;	Mismatches 151;	Indels 27; Gaps 5;
QY 524	GGATTTACTTACTGAACCCCATCTGCTCCCTGACCGCAT	TGAGCTTGAGGCCAACATCAATG	583	
DL 563	GGGTAAACTGGATGAACCCCTNGTATGGGAGAGAATNGAGNACACCNATTTCTTGGG		504	

Query Match 5.4%; Score 58.4; DB 4; Length: 7902;  
Best Local Similarity 52.5%; Pred. No. 2.3e-08;  
Matches 128; Conservative 0; Mismatches 116; Indels 0; Gaps 0;

Db 4961 TGAAGGCAATTTATCTTAGCA3CGGGATTGGGAACCTCGCTTGGCTCTATGACTGAAATA 5020  
Qy 62 TCCGAAACCCCTCGTGGATTTCGAAACAGCCCATGATTTGACACAGATCGAGCTT 121  
Db 5021 CCCCTAAAGCCTTGGTTTCAG3TTAATCAAAACCTTTGATTGAGTACCAAAATTGAGTTTC 5080  
Qy 122 TGAAGAAGTTGGGGTCACAGAGTGGTTTGGCTATCAACATCTCGCCCGAGGTAATGA 181  
Db 5081 TCAAGAAAAGGAATCAAT3CATCATCATCTGTTGGTTATCTTTAAAGAACAAATCG 5140  
Qy 182 TTAATTTCTTGAAGGACTTTTGAAGTAAGCTTTGGCATCAACAATPACATGCTCCCAAGAGA 241  
Db 5141 ATTACTTGAAGAGAAATAC3GTGTCGTCCTCGTTTCAATGATAAATACGCTGACTACA 5200  
Qy 242 CTGA 245  
Db 5201 ATAA 5204

## RESULT 5

US-09-024-023-1  
; Sequence 1, Application US/09024023  
; Patent No. 6110899  
; GENERAL INFORMATION:  
; APPLICANT: Lonetto, Michael A.  
; TITLE OF INVENTION: l1cc  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dechert, Price & Rhoads  
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre  
; CITY: Philadelphia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19103-2793  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/024,023  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/039,210  
; FILING DATE: 28-FEB-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Falk, Stephen T  
; REGISTRATION NUMBER: 36,795  
; REFERENCE/DOCKET NUMBER: GM50018  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 215-994-2488  
; TELEFAX: 215-994-2222  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 690 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear

## US-09-024-023-1

Query Match 5.1%; Score 55.2; DB 3; Length 690;  
Best Local Similarity 51.6%; Pred. No. 6.1e-08;  
Matches 126; Conservative 0; Mismatches 118; Indels 0; Gaps 0;  
Qy 2 TGAAGGCCCTCATTTCTGCGGGGTTTCGGAACCCGCTTCGGCCTTTGACTCTGAGCT 61  
Db 2 TGAAGGCCATTATCTTAGCAGCGGGAATGGGAACCTCGCTTCGCTCTATGACTGAAATA 61  
Qy 62 TCCGAAACCCCTCGTGGATTTCGAAACAGCCCATGATTCGACACAGATCGAGCTT 121  
Db 62 CCCCTAAAGCCTTGGTTTCAGGTTAATCAAAACCTTTGATTGAATACCAAAATTGAGTTTC 121

Qy 122 TGAAGAAGTTGGGGTCACAGAGTGGTTTGGCTATCAACTATCCCCAGAGGTAATGA 181  
Db 122 TCAAGAAAAGGAATCAAT3CATCATCATCATCATCGTTGGTTATCTTTAAAGAACAAATTCG 181  
Qy 182 TTAATTTCTTGAAGGACTTTTGAAGTAAGCTTTGGCATCAACAATPACATGCTCCCAAGAGA 241  
Db 182 ATATTAAAGAAAATACG3GTGTCGCTCGTTTCAATGATAAATACGCTGACTACA 241  
Qy 242 CTGA 245  
Db 242 ATAA 245

## RESULT 6

US-09-531-111-1  
; Sequence 1, Application US/09531111  
; Patent No. 6303571  
; GENERAL INFORMATION:  
; APPLICANT: Lonetto, Michael A.  
; TITLE OF INVENTION: l1cc  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dechert, Price & Rhoads  
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre  
; CITY: Philadelphia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19103-2793  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/531,111  
; FILING DATE: 17-Mar-2000  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/024,023  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Falk, Stephen T  
; REGISTRATION NUMBER: 36,795  
; REFERENCE/DOCKET NUMBER: GM50018  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 215-994-2488  
; TELEFAX: 215-994-2222  
; TELEX: <Unknown>  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 690 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-531-111-1

Query Match 5.1%; Score 55.2; DB 4; Length 690;  
Best Local Similarity 51.6%; Pred. No. 6.1e-08;  
Matches 126; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

Qy 2 TGAAGGCCCTCATTTCTGCGGGGTTTCGGAACCCGCTTCGGCCTTTGACTCTGAGCT 61  
Db 2 TGAAGGCCATTATCTTAGCAGCGGGAATGGGAACCTCGCTTCGCTCTATGACTGAAATA 61  
Qy 62 TCCGAAACCCCTCGTGGATTTCGAAACAGCCCATGATTCGACACAGATCGAGCTT 121  
Db 62 CCCCTAAAGCCTTGGTTTCAGGTTAATCAAAACCTTTGATTGAATACCAAAATTGAGTTTC 121  
Qy 122 TGAAGAAGTTGGGGTCACAGAGTGGTTTGGCTATCAACTATCGCCCGAGGTAATGA 181  
Db 122 TCAAGAAAAGGAATCAAT3CATCATCATCATCATCGTTGGTTATCTTTAAAGAACAAATTCG 181

Qy 182 TTAATTTCTGAAGGACTTTGAGGATAAGCTTGGCATCACAATTACATGCTCCCAAGAGA 241  
| | | | |  
Db 182 ATTATTTAAAGAAAAATACGGTGTTCGCTCGTTTTCATGATAAATACGCTGACTACA 241  
| | | | |  
Qy 242 CTGA 245  
| | | | |  
Db 242 ATAA 245

RESULT 7  
US-09-024-023-3  
; Sequence 3, Application US/09024023  
; Patent No. 6110899  
; GENERAL INFORMATION:  
; APPLICANT: Lonetto, Michael A.  
; TITLE OF INVENTION: licc  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dechert, Price & Rhoads  
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre  
; CITY: Philadelphia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19103-2793  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Fast-SEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/024,023  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/035,210  
; FILING DATE: 28-FEB-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Falk, Stephen T  
; REGISTRATION NUMBER: 36,795  
; REFERENCE/DOCKET NUMBER: GM50018  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 215-994-2488  
; TELEFAX: 215-994-2222  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 287 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear

US-09-024-023-3  
Query Match 4.9%; Score 53.6; DB 3; Length 287;  
Best Local Similarity 51.2%; Pred. No. 1.2e-07;  
Matches 125; Conservative 0; Mismatches 119; Indels 0; Gaps 0;  
Qy 2 TGAAGGCCCTCATTTCTTGTGCGGGGTTTCGGAACCCGCTTTCGGCCTTTGACTCTGAGCT 61  
| | | | |  
Db 2 TGAAGGCCATTATCTTAGCGCGGATTGGGAACCTCGCTTGCCTATGACTGAAATA 61  
| | | | |  
Qy 62 TCCCAAAACCCCTCGTGGATTTTGCAGAACCCCATGATTCGACACAGATCGAGCTT 121  
| | | | |  
Db 62 CCCCTAAAGCCCTTGGTTACAGGTTAATCAAAAACCTTTGATTGAATACCAAAATTCGATTC 121  
| | | | |  
Qy 122 TCAAGAGAGTTGGGTGCAGAGGTGTTTTCGCTATCACTATCGCCACAGATCAATGA 181  
| | | | |  
Db 122 TCAAGAAAAAGGAATCAATGACATCATCATCGTTGGTTTATCTTAAAGAACAAATTCG 181  
| | | | |  
Qy 182 TTAATTTCTGAAGGACTTTGAGGATAAGCTTGGCATCACAATTACATGCTCCCAAGAGA 241  
| | | | |  
Db 182 ATTATTTAAAGAAAAATACGGTGTTCGCTCGTTTTCATGATAAATACGCTGACTACA 241  
| | | | |

Qy 242 CTGA 245  
| | | | |  
Db 242 ATAA 245  
| | | | |  
RESULT 8  
US-09-531-111-3  
; Sequence 3, Application US/09531111  
; Patent No. 6303571  
; GENERAL INFORMATION:  
; APPLICANT: Lonetto, Michael A.  
; TITLE OF INVENTION: licc  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dechert, Price & Rhoads  
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre  
; CITY: Philadelphia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19103-2793  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Fast-SEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/531,111  
; FILING DATE: 17-Mar-2000  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/024,023  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Falk, Stephen T  
; REGISTRATION NUMBER: 36,795  
; REFERENCE/DOCKET NUMBER: GM50018  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 215-994-2488  
; TELEFAX: 215-994-2222  
; TELEX: <Unknown>  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 287 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:  
US-09-531-111-3

Query Match 4.9%; Score 53.6; DB 4; Length 287;  
Best Local Similarity 51.2%; Pred. No. 1.2e-07;  
Matches 125; Conservative 0; Mismatches 119; Indels 0; Gaps 0;  
Qy 2 TGAAGGCCCTCATTTCTTGTGCGGGGTTTCGGAACCCGCTTTCGGCCTTTGACTCTGAGCT 61  
| | | | |  
Db 2 TGAAGGCCATTATCTTAGCAGCGGATTGGGAACCTCGCTTGCCTATGACTGAAATA 61  
| | | | |  
Qy 62 TCCCAAAACCCCTCGTGGATTTTGCAGAACCCCATGATTCGACACAGATCGAGCTT 121  
| | | | |  
Db 62 CCCCTAAAGCCCTTGGTTACAGGTTAATCAAAAACCTTTGATTGAATACCAAAATTCGATTC 121  
| | | | |  
Qy 122 TGAAGAGAGTTGGGTGCAGAGGTGTTTTCGCTATCACTATCGCCACAGAGCTAATGA 181  
| | | | |  
Db 122 TCAAGAAAAAGGAATCAATGACATCATCATCGTTGGTTTATCTTAAAGAACAAATTCG 181  
| | | | |  
Qy 182 TTAATTTCTGAAGGACTTTGAGGATAAGCTTGGCATCACAATTACATGCTCCCAAGAGA 241  
| | | | |  
Db 182 ATTATTTAAAGAAAAATACGGTGTTCGCTCGTTTTCATGATAAATACGCTGACTACA 241  
| | | | |  
Qy 242 CTGA 245  
| | | | |  
Db 242 ATAA 245

RESULT 9  
US-08-196-218-31  
; Sequence 31, Application US/08196218  
; Patent No. 5614619  
; GENERAL INFORMATION:  
; APPLICANT: Piepersberg, Wolfgang  
; APPLICANT: Stockmann, Michael  
; APPLICANT: Taleghani, Kampiz Mansouri  
; APPLICANT: Distler, Jurgen  
; APPLICANT: Grabley, Susanne  
; APPLICANT: Sichel, Petra  
; APPLICANT: Brau, Barbara  
; TITLE OF INVENTION: Secondary-Metabolite Biosynthesis Genes  
; TITLE OF INVENTION: From Actinomycetes, Method of Isolating Them, and Their  
; TITLE OF INVENTION: Use.  
; NUMBER OF SEQUENCES: 34  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
; STREET: 1300 I Street, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: United States  
; ZIP: 20005-3315  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/196,218  
; FILING DATE: 25-AUG-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ogden, Stasia L.  
; REGISTRATION NUMBER: 36,223  
; REFERENCE/DOCKET NUMBER: 03481.1372-00000  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-408-4000  
; TELEFAX: 202-408-4400  
; INFORMATION FOR SEQ ID NO: 31:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2634 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 3..401  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 416..1531  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1561..2625  
US-08-196-218-31  
  
Query Match 4.9%; Score 52.8; DB 1; Length 2634;  
Best Local Similarity 57.1%; Pred. No. 8.4e-07;  
Matches 96; Conservative 0; Mismatches 72; Indels 0; Gaps 0;  
  
Qy 1 ATGAAGGCCCTCATCTTGTGGGGTTTCGAAACCCGCTTCGGCCTTTGACTCTGAGC 60  
Db 1561 ATGAAGGCTCTGGTCTGCGCGGATCTCGTACCCGCTTCGGCCTTTTTCAGTTATTCG 1620  
Qy 61 TTCCCGAAACCCCTCGTGGATTTTGCAACAGCCCATGATCTGCACACAGATCGAAGCT 120  
Db 1621 ATGCCAAACAACTGATCCCATCCCAACACACCCGCTGCTGGTCATGCTGCTGAAGCC 1680  
Qy 121 TTGAAGAAGTTGGGTCACAGAGTGTTTTTGGCTATCAACTATCGC 168

Db 1681 GTCCGGAGCTGGCGTGCACCGAGGTGCGGTCATCGTCGCAACCGC 1728  
  
RESULT 10  
US-08-681-953-31  
; Sequence 31, Application US/08681953  
; Patent No. 5710032  
; GENERAL INFORMATION:  
; APPLICANT: Piepersberg, Wolfgang  
; APPLICANT: Stockmann, Michael  
; APPLICANT: Taleghani, Kampiz Mansouri  
; APPLICANT: Distler, Jurgen  
; APPLICANT: Grabley, Susanne  
; APPLICANT: Sichel, Petra  
; APPLICANT: Brau, Barbara  
; TITLE OF INVENTION: Secondary-Metabolite Biosynthesis Genes  
; TITLE OF INVENTION: From Actinomycetes, Method of Isolating Them, and Their  
; TITLE OF INVENTION: Use.  
; NUMBER OF SEQUENCES: 34  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
; STREET: 1300 I Street, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: United States  
; ZIP: 20005-3315  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/681,953  
; FILING DATE: 30-JUL-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/196,218  
; FILING DATE: 25-AUG-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ogden, Stasia L.  
; REGISTRATION NUMBER: 36,228  
; REFERENCE/DOCKET NUMBER: 02481.1372-00000  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-408-4000  
; TELEFAX: 202-408-4400  
; INFORMATION FOR SEQ ID NO: 31:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2634 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 3..401  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 416..1531  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1561..2625  
US-08-681-953-31  
  
Query Match 4.9%; Score 52.8; DB 1; Length 2634;  
Best Local Similarity 57.1%; Pred. No. 8.4e-07;  
Matches 96; Conservative 0; Mismatches 72; Indels 0; Gaps 0;  
  
Qy 1 ATGAAGGCCCTCATCTTGTGGGGTTTCGAAACCCGCTTCGGCCTTTGACTCTGAGC 60  
Db 1561 ATGAAGGCTCTGGTCTGCGCGGATCTCGTACCCGCTTCGGCCTTTTTCAGTTATTCG 1620  
Qy 61 TTCCCGAAACCCCTCGTGGATTTTGCAACAGCCCATGATCTGCACACAGATCGAAGCT 120

Db 1621 ATGCCCAAACTGATCCCATGCCAACACACCCGGTGGTGTGCTGAACGCC 1683  
QY 121 TTGAAGAAGTTGGGTCACAGAGTGTTTGGTATCAACTATCGC 168  
Db 1691 GTCCGGAGCTGGGGCTGACAGGTGGCGTCATCGTCGCAACGCC 1728

## RESULT 11

US-09-370-838-155  
; Sequence 155, Application US/09370838  
; Patent No. 6444425  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Mohamath, Roadoh  
; APPLICANT: Secrist, Heather  
; TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS OF  
; TITLE OF INVENTION: LUNG CANCER AND METHODS FOR THEIR USE  
; FILE REFERENCE: 210121.475C1  
; CURRENT APPLICATION NUMBER: US/09/370,838  
; CURRENT FILING DATE: 1999-08-09  
; EARLIER APPLICATION NUMBER: US 09/285,323  
; EARLIER FILING DATE: 1999-04-02  
; NUMBER OF SEQ ID NOS: 289  
; SOFTWARE: FastSeq for Windows Version: 3.0  
; SEQ ID NO 155  
; LENGTH: 678  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-370-838-155

Query Match 4.3%; Score 46.2; DB 4; Length 678;

Best Local Similarity 49.5%; Pred. No. 5.5e-05;  
Matches 151; Conservative 0; Mismatches 148; Indels 6; Gaps 1;

- QY 235 CAAGAGACTGAGCCCTTAGGACCCCTGCTCTAGCAAGGCAAGCTGCG 294  
Db 350 CAGGAATTGGCCCTTAGGCACAGGGGTGCTTTACCATTTTCGAGACCATCTG 409  
QY 295 GATGATCTGGCCAGCCATTCTTTGCTCTCAACAGTGTATGAAGCGAATACCCATT 354  
Db 410 GCTGGAGCCCGAGGCACTTCTGCTCAATGCTGATGCTGCTCGACTTCCCTTG 469  
QY 355 GCTGAACATCAAAATTCACAGTGTATGTTGTGTGAGGCAAAATATGTCACTAAG 414  
Db 470 AGTGTATGTTGGAAGCCCPACGACGCGTCAACCCCTTTCTACTCCTTGGCACTACG 529  
QY 415 GTGGATG-----AACCATCAAAATACGCTGTTGTGTTATGGAGGAGGCAACTGGCAGG 468  
Db 530 GCTAACAGGACCATCCCTCAACTACGCTGTCATCGTTGAGATCCACAGACACAG 589  
QY 469 GTGGAAGGTTTGTGAGAGCCAAAATATTGTGGGTAAACAAGATCAATGCTGGGATT 528  
Db 590 GTATTGCACTATGTGGAGAACCCAGCACATTTATCAGTGACATCATCAACTGGGCACC 649  
QY 529 TACTT 533  
Db 650 TACCT 654

## RESULT 12

US-09-107-532A-2711  
; Sequence 2711, Application US/09107532A  
; Patent No. 6583275  
; GENERAL INFORMATION:  
; APPLICANT: Lynn A Doucette-Stamm and David Bush  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; NUMBER OF SEQUENCES: 7310  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
; STREET: 100 Beaver Street

CITY: Waltham  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02354  
COMPUTER READABLE FORM:  
MEDIUM TYPE: CD-ROM ISO9660  
COMPUTER: PC  
OPERATING SYSTEM: <Unknown>  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/107,532A  
FILING DATE: 30-Jun-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/085,598  
FILING DATE: 14 May 1998  
APPLICATION NUMBER: 60/051571  
FILING DATE: July 2, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Ariniello, Pamela Deneke  
REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-012  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781)893-5007  
TELEFAX: (781)893-8277  
INFORMATION FOR SEQ ID NO: 2711:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 891 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: circular  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Enterococcus faecium  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (8) LOCATION 1...891  
SEQUENCE DESCRIPTION: SEQ ID NO: 2711:  
US-09-107-532A-2711

Query Match 3.9%; Score 42.2; DB 4; Length 891;

Best Local Similarity 55.9%; Pred. No. 0.0013;  
Matches 80; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 1 ATGAAGGCCCTCATCTCTCGGGGTTTCGAAACCCGCTTCGGCCTTTGACTCTGAGC 60  
Db 25 ATGAAGGGAATCATCTCTCGGGGCGAGCGCAACACGCTCTGTACCCCTTAAACAAAGCA 84  
QY 61 TTCCCGAAACCCCTCGTGGATTTTCAACACAGCCCATGATCTGCACCATCGAAGCT 120  
Db 85 ACATCGAAACAATGATGCCGATTTATGACAAACCAATGATTTATTCCAATGCTACT 144  
QY 121 TTGAAGAAGTTGGGCTCACAGA 143  
Db 145 CTGATGTTGGCGGAATCAAAGA 167

## RESULT 13

US-09-328-352-3686  
; Sequence 3686, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: GTC99-03PA  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 9252  
; SEQ ID NO 3686  
; LENGTH: 702  
; TYPE: DNA



ORGANISM: Acinetobacter baumannii  
US-09-328-352-3686

Query Match 3.8%; Score 41.8; DB 4; Length 702;  
Best Local Similarity 50.8%; Pred. No. 0.0016;  
Matches 100; Conservative 0; Mismatches 97; Indels 0; Gaps 0;

QY 1 ATGAAGGCCCTCATCTTGTGCGGGGTTTCGGAACCGCCTTGGCGCTTTGACTCTGAGC 60  
DB 13 ATGAAGCAATGATTTTAGCTGTGCTGGCTTGGCAATCGTATGGTCCACTACACTATAT 72  
CY 61 TTCCCGAAGCCCTCGTGGGATTTTGCACCAAGCCCATGATTTGCACCCAGACGCGAAGCT 120  
DB 73 ACACCAAGCCTCTGCTTGAAGTAGGGGTAAAGCACTATTGTGGCATATTGAATA 132  
QY 121 TTGAAGAGAGTTGGGTCACAGAGTGGTTTTTGGCTATCAACTATATGCCCGCAGAGGTAATG 180  
DB 133 CTTAAGAAATCGGTGTGACGAAATAGTCATCAATCTCGGTGGTTAGCAGACAACTC 192  
QY 181 ATTAATTTCTGAAGGA 197  
DB 193 ATCAGTAGCTGGGAGA 209

RESULT 14  
US-09-194-905-7  
; Sequence 7, Application US/09194905  
; Patent No. 630627  
; GENERAL INFORMATION:  
; APPLICANT: DECKER, Heinrich  
; TITLE OF INVENTION: ISOLATION OF THE BIOSYNTHESIS GENES FOR  
; TITLE OF INVENTION: PSEUDO-OLIGOSACCHARIDES FROM STREPTOMYCES GLAUCESCENS  
; TITLE OF INVENTION: GLA.O AND THEIR USE  
; NUMBER OF SEQUENCES: 13  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FOLEY & LARDNER  
; STREET: 3000 K Street, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/M3-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; APPLICATION NUMBER: US/09/194,905  
; FILING DATE: 29-JUL-1998  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: WO PCT/EP97/02826  
; FILING DATE: 30-MAY-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: DE 19622783.6  
; FILING DATE: 07-JUN-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Granados, Patricia D.  
; REGISTRATION NUMBER: 33,683  
; REFERENCE/DOCKET NUMBER: 026083/0193  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 672-5300  
; TELEFAX: (202) 672-5399  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 6854 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
US-09-194-905-7

Query Match 3.8%; Score 41.6; DB 4; Length 6854;

Best Local Similarity 55.6%; Pred. No. 0.0071;  
Matches 80; Conservative 0; Mismatches 64; Indels 0; Gaps 0;  
QY 2 TGAAGGCCCTCATCTTGTGCGGGGTTTCGGAACCGCCTTGGCGCTTTGACTCTGAGCT 61  
DB 2269 TGAAGGCCCTGTGCTGCGAGGTGGAACCGGCGAGCAGACTGAGGCGCTTCAACCCACCG 2328  
QY 62 TCCGAAACCCCTCGTGGATTTTGCACCAAGCCCATGATTTGCACCGATCGAAGCTT 121  
DB 2329 CGCCCAAGCAGCTGCTCCCACTCGCCCAACAGCCCGTCTCTTCTACGCGTGGAGTCCC 2388  
QY 122 TGAAGAAGTTCGGGTCAACAGG 145  
DB 2389 TCGCCGCGCGGTGTCGGGAGG 242

RESULT 15  
US-09-221-017B-904  
; Sequence 904, Application US/09221017B  
; Patent No. 6444799  
; GENERAL INFORMATION:  
; APPLICANT: ROSS, Bruce C.  
; TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF  
; NUMBER OF SEQUENCES: 1120  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORRISON & FORSTER  
; STREET: 755 PAGE MILL ROAD  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304-1018  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows  
; SOFTWARE: FastSEQ for Windows Version 2.0b  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/221,017B  
; FILING DATE: 23-DEC-1998  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PP1182  
; FILING DATE: 31-DEC-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PP1546  
; FILING DATE: 30-JAN-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PP2911  
; FILING DATE: 09-APR-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/AU98/01023  
; FILING DATE: 10-DEC-1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Monroy, Gladys H  
; REGISTRATION NUMBER: 32,430  
; REFERENCE/DOCKET NUMBER: 27340-20021.00  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650-813-5600  
; TELEFAX: 650-494-0792  
; TELEX: 706141  
; INFORMATION FOR SEQ ID NO: 904:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3571 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHETICAL: NO  
; ANTI-SENSE: UNKNOWN  
; ORIGINAL SOURCE:  
; ORGANISM: PORPHYROMONAS GINGIVALIS  
; FEATURE:  
; NAME/KEY: misc\_feature

Search completed: October 14, 2003, 14:18:50  
Job time : 237 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 15, 2003, 10:53:48 ; Search time 362 Seconds

(without alignments)  
7784.434 Million cell updates/sec

Title: US-09-374-967-1

Perfect score: 1086

Sequence: 1 atgaagccctcattcttgt.....agctgagatcgctcatgtga 1086

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1731049 seqs, 1297405648 residues

Total number of hits satisfying chosen parameters: 3462098

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:\*

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
16: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*  
17: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	690.8	63.6	1109	12	US-10-342-224-23
2	399.6	36.8	1559	12	US-10-427-631-17
3	360.8	33.2	1095	14	US-10-128-714-7009
4	310.8	28.6	1089	12	US-10-032-585-6223
5	247.6	22.8	1051	14	US-10-128-714-1009
6	247.6	22.8	1370	14	US-10-128-714-6009
7	247.6	22.8	3051	14	US-10-128-714-9
8	247.6	22.8	3370	14	US-10-128-714-5009
9	241.6	22.2	960	14	US-10-128-714-2009
10	187	17.2	212	11	US-09-294-093B-2164
11	180.8	16.6	488	11	US-09-770-961-719
12	151.4	13.9	364	10	US-09-878-574-2725
13	130.6	12.0	269	10	US-09-878-574-8673
14	79.2	7.3	272	10	US-09-878-574-10462
15	74.6	6.9	2493	14	US-10-156-761-6955
16	74.6	6.9	9025608	14	US-10-156-761-1

c	17	73.2	6.7	787	10	US-09-954-773A-20	Sequence 20, Appl
	18	64.2	5.9	456	10	US-09-796-692-5169	Sequence 5169, Ap
	19	64.2	5.9	456	14	US-10-040-862-5169	Sequence 5169, Ap
	20	55.2	5.1	282	9	US-09-294-093B-3871	Sequence 3871, Ap
	21	50.8	4.7	1080	14	US-10-156-761-5021	Sequence 5021, Ap
	22	46.2	4.3	678	10	US-09-738-973-155	Sequence 155, App
	23	46.2	4.3	678	10	US-09-854-133-155	Sequence 155, App
	24	46.2	4.3	678	14	US-10-144-649A-155	Sequence 155, App
	25	42.8	3.9	6158	12	US-10-292-198-1	Sequence 1, Appli
	26	41.6	3.8	6854	10	US-09-922-683-7	Sequence 7, Appli
	27	41.2	3.8	386	11	US-09-918-995-7323	Sequence 7323, Ap
	28	39	3.6	474	11	US-09-918-995-24581	Sequence 24581, A
c	29	38	3.5	167343	9	US-09-962-436-281	Sequence 281, App
c	30	38	3.5	167343	10	US-09-964-824A-273	Sequence 273, App
c	31	37.8	3.5	897	14	US-10-156-761-935	Sequence 935, App
c	32	37.8	3.5	9025608	14	US-10-156-761-1	Sequence 1, Appli
	33	37.6	3.5	2840	12	US-10-114-153-3	Sequence 3, Appli
	34	37.4	3.4	2115	14	US-10-128-714-7238	Sequence 7238, Ap
	35	37.4	3.4	5144	14	US-10-043-344-105	Sequence 105, App
	36	36.8	3.4	536165	11	US-09-939-964-1	Sequence 1, Appli
c	37	36.6	3.4	1179	14	US-10-156-761-7119	Sequence 7119, Ap
c	38	36.4	3.4	855	10	US-09-738-626-371	Sequence 371, App
c	39	36.4	3.4	3309400	10	US-09-738-626-1	Sequence 1, Appli
	40	36	3.3	317	14	US-10-146-938-10	Sequence 10, Appl
	41	35.8	3.3	4699	14	US-10-043-344-1	Sequence 1, Appli
c	42	35.6	3.3	1830121	14	US-10-329-960-1	Sequence 1, Appli
c	43	35.4	3.3	3119	12	US-09-975-719-173	Sequence 173, App
c	44	35.2	3.2	808	10	US-09-954-773A-19	Sequence 19, Appl
	45	35.2	3.2	879	10	US-09-861-289-11	Sequence 11, Appl

#### ALIGNMENTS

#### RESULT 1

US-10-342-224-23  
; Sequence 23, Application US/10342224  
; Publication No. US20030162294A1  
; GENERAL INFORMATION:  
; APPLICANT: Nathalie Verbruggen  
; TITLE OF INVENTION: Genes Involved in Tolerance to Environmental Stress  
; FILE REFERENCE: CN-012US  
; CURRENT APPLICATION NUMBER: US/10/342,224  
; PRIOR FILING DATE: 2003-03-13  
; PRIOR APPLICATION NUMBER: US/09/762,154  
; PRIOR FILING DATE: 2002-02-02  
; PRIOR APPLICATION NUMBER: EP 98202634.6  
; PRIOR FILING DATE: 1998-08-04  
; NUMBER OF SEQ ID NOS: 123  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 23  
; LENGTH: 1109  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (22)....(1107)  
US-10-342-224-23

Query Match 63.6%; Score 690.8; DB 12; Length 1109;  
Best Local Similarity 77.3%; Pred. No. 6.7e-220;  
Matches 839; Conservative 0; Mismatches 247; Indels 0; Gaps 0;

QY	1	ATGAGGCCCTCATTTCTGTCGGGGGTTTCGAAACCCGCTTCGGCCTTGACTCTGAGC	60
DB	22	ATGAGGCCCTCATTTCTGTCGGGGGTTTCGAAACCCGCTTCGGCCTTGACTCTGAGC	81
QY	61	TTCCCGAAACCCCTCGTGGATTTCGAAACAGCCATGATTCGACCATCGAAGCT	120
DB	82	TTCCCGAAACCCCTCGTGGATTTCGAAACAGCCATGATTCGACCATCGAAGCT	141
QY	121	TTGAAAGAGTGGGGTCACAGAGTGGTGGTTCGATCAACTGCGCCAGAGTAATG	180

Db 142 CTTAAGCAGTTGGAGTTGATGAAGTGGTTTGGCCATCAATTATCAGCCAGAGGTGATG 201  
Qy 181 ATTAATTTCTGAAGGACTTTAGAGTAAGCTTTGGCATCAAAATTACATGCTCCCAAGAG 240  
Db 202 CTGAACCTTTCTGAAGGACTTTAGACCAAGCTGGAAATCAAAATCACTTGTCTCACAAGAG 261  
Qy 241 ACTGAGCCCTTAGGAACCGCTGGCCCTCTTCTCTAGCAAGGACAGCTTGGCGATGA 300  
Db 262 ACCGAGCCACTAGTAGCCGCTGGCTCTGCTCTAGCGAGACAAATTTGCTTGATGA 321  
Qy 301 TCTGGCAGCCATTCTTTGTCTCAACAGTGTATCTCATAGCGAATACCCATTGTCTGAA 360  
Db 322 TCTGGAGGCCCTTTCTTTGTCTTAAAGTGTATAGTGTAGTACCTCTTAAAGAA 381  
Qy 361 CTATCAAAATTTCAAGTGTATGCTGGTGGAGCAAAATTTATGTCATCAAGGTGGAT 420  
Db 382 ATGCTTGAGTTTCAAAATCTCAGCGTGGGAGAGCCCTCCATATGTTGTAACAAAGGTGGAT 441  
Qy 421 GAACCATCAAAATACGGTGTGTGTTTATCGGAGGAGCACTGCGAGGTTGGAAGGTTT 480  
Db 442 GAACCGTGAATATGAGTGTGTTTATGGAAGAAAGCACTGGAAGAGTGGAGAAATTT 501  
Qy 481 GTTGAGAGCCAAATAATTTGTGGGTAAACAAGATCAATGCTGGGATTTACTTACTGAAC 540  
Db 502 GTGGAAGCCAAATCTATGTAGGTAAACAAGATCAACGCTGGGATTTATCTTCTGAAC 561  
Qy 541 CCATCTGCTTGAACCGATTTAGCTGAGGCAACATCAATTTGAGAAAGAGGCTTCCCT 600  
Db 562 CCATCTGCTTGAATAAGATTAGCTAAGACCGACTTCAATCGAAAGAGAGACTTCCCT 621  
Qy 601 CAATTCGAGCTGATCAACAACCTATGCAATGCTCTCCAGTGTTCAGGATTTTGGATGGATTT 660  
Db 622 AAGATTTGAGCGGCAAGGCTCTATGCTATGCTGTACACGGGTTTGGATGGACATTT 681  
Qy 661 GGTGAGCTAGGCACTACATTTACTGCTGCTGCTTTATCTAGACTCGATTTAGAGAA 720  
Db 682 GGGCAACCCCGTGACTACATTAACGGGTTTGAGACTCTTACTTACTTCCCTTAGAGAAA 741  
Qy 721 TCAGTCCCAAGCTAGCTACTGGAGCAATGTTGTTGGCAATGTGCTGGTGCATGAGAGC 780  
Db 742 TCTCTGCCAAATTAACCAAGTGGGCCACACATAGTTGGGAATGTTCTTTGTGACGAACC 801  
Qy 781 GCCAAGATTGGAGAGGTGTCTGATTTGCTCTCATGTCGCCATTGGACCTGGGTGTT 840  
Db 802 GCTACAAATGGGGAAGATGTTGATTTGACCAAGCTTGCATTTGTCTCCAGGCTGCATTT 861  
Qy 841 GTGAGGACGGCTGAGGCTTTCCCGCTGCACTGTCAATGCGCGCGTGCCTATCAAGAG 900  
Db 862 GTTGAGTCAGGAGTCAGACTCTCCCGATGCACGGTCAATGCTGGAGTCGCATCAAGAG 921  
Qy 901 CATGCTTGATCTCAACAGCAATTTATCGGCTGGGCACTCAACTGTTGTGATCAATGGGACGG 960  
Db 922 CATCGGTGTATCTCGAGCAGPATCATCGGTTGGGCACTCAACGGTTGTGATCAATGGGCCAGG 981  
Qy 961 ATAGAGATATGATATCTCTCGGAGAGTGTTCATGTTGTGATGAGGTGTACAGCAAT 1020  
Db 982 ATCGAGAACATGACATCTCTCGTGAAGATGTTTCATGTTGAGCGATGATGATCTATAGCAAT 1041  
Qy 1021 GCGCGTGTGTTCTCCACAATAAGAGATCAAGTCAAGCAATTTGAAAGCCTGAGATCGTC 1080  
Db 1042 GGAGGAGTTGTTTGGCCACAACAGGAGATCAAAATCAAAACATCTTTGAAGCCAGAGATAGT 1101  
Qy 1081 ATGTGA 1086  
Db 1102 ATGTGA 1107

## RESULT 2

US-10-427-631-17

; Sequence 17, Application US/10427631

; Publication No. US20030175923A1

; GENERAL INFORMATION:

; APPLICANT: INCYTE CORPORATION; TANG, Y. Tom;

; APPLICANT: CORLEY, Neil C.; GUESLER, Karl J.;  
; APPLICANT: BAUGHN, Mariah R.; LAL, Preeti G.;  
; APPLICANT: YUE, Henry; HILLMAN, Jennifer L.;  
; APPLICANT: AZIMZAI, Yalda  
; TITLE OF INVENTION: HUMAN TRANSFERASE PROTEINS  
; FILE REFERENCE: PF-0592-1 DIV  
; CURRENT APPLICATION NUMBER: US/10/427,631  
; CURRENT FILING DATE: 2003-04-29  
; PRIOR APPLICATION NUMBER: US 09/786,240  
; PRIOR FILING DATE: 2002-03-12  
; PRIOR APPLICATION NUMBER: PCT/US99/20989  
; PRIOR FILING DATE: 1998-09-09  
; PRIOR APPLICATION NUMBER: US 60/172,220  
; PRIOR FILING DATE: 1998-09-10  
; PRIOR APPLICATION NUMBER: US 60/155,248  
; PRIOR FILING DATE: 1998-11-04  
; PRIOR APPLICATION NUMBER: US 60/133,642  
; PRIOR FILING DATE: 1998-05-11  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: PERU Program  
; SEQ ID NO 17  
; LENGTH: 1559  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: Incyte ID No. US20030175923A1 2682663CB1  
US-10-427-631-17

Query Match 36.8%; Score 399.6; DB 12; Length 1559;  
Best Local Similarity 61.1%; Pred. No. 2.2e-122; Indels 3; Gaps 1;  
Matches 664; Conservative 0; Mismatches 419;  
Qy 1 ATGAAGGCCCTCAATCTTGTGCGGGGTTTTCGAAACCCGCTTTCGGCCTTTGACTCTGAGC 60  
Db 213 ATGAAGGCACTGATCTTAGTGGGGGCTATGGACGCGCTACGCGCGTGAGCTGAGC 272  
Qy 61 TTCCCGAAACCCCTCGTGGATTTTCAACAACAGCCCATGATCTCGACCAAGATCAAGCT 120  
Db 273 ACCCCGAAGCCACTGCTGGACTCTGCAATAGCCCATCTTGTGCACCAAGTGAGGCG 332  
Qy 121 TTGAAGAAGTTGGGTGCACAGAGTGGTTTGGTCTATCAACTATCGCCACAGGTAATG 180  
Db 333 CTAGCGCGGAGGCGTGAGCCACGTGATCTTGGCGGTGAGCTATGCTCGCAGGTCGTG 392  
Qy 181 ATTAATTTCTGAAGGACTTTTGAGGATAAGCTTGGCATCACAATTTACATGCTCCCAAGAG 240  
Db 393 GAGAAGGAAATGAAGGCACAGGACAGAGGCTGGGAATCCGAATCTCCATGTCCCATGAA 452  
Qy 241 ACTGAGCCCTTAGGAACCGCTGGCCCTTGTGCTTAGCAAGGACAAAGCTTGGCGATGA 300  
Db 453 GAGGAGCCCTTTGGGACACACTGGGCCCTTGGCGCTGGGCCCTGACTCTCTC---TGAG 509  
Qy 301 TCTGCGCAGCCATTCTTTGCTCAACAGTGTATGTCATAAGCGAATACCAATTTGCTGAA 360  
Db 510 ACTGACAGCCCTTTCTTGTCTCTCAACAGTACGTCATCTGCGATTTCCCTTCCAAACC 569  
Qy 361 CTCATCAAAATTTCAAGTGTCTATGGTGGAGCAAAATTTATGTCATTAAGTGTGAT 420  
Db 570 ATGGTGCACTTCCACCGGCACCATGCCAGGAGGCTCCATCCTGTCACCAAGGTGAG 629  
Qy 421 GAACCATCAAAATACGGTGTGTTGTTATGAGGAGGCAACTGGCAGGCGTGGAAAGGTTT 480  
Db 630 GAACCCCTCCAGTACGGTGTGTTGTTGAGGCTGACACAGGCCCGCATTCACCGGTTTC 689  
Qy 481 GTTGAGAGCCAAATAATTTGTGGTAAACAGATCAATGCTGGGATTTTACTTACTGAAC 540  
Db 690 GTGGAGAGCCACAGGTGTTGTGTCCAAATAGATCAACGAGGCGATGTACATCTGAGC 749  
Qy 541 CCATCTGCTTTCACCGCATTTGAGCTGAGGCGCAACATCAATTTGAGAAAGAGGCTTCCCT 600  
Db 750 CTGCGAGTGTCTCCGGCGCATCCAGCTGGCAGCTAGTCCA-TTGAGAAGAGAGGCTCTCCCC 809



QY 1069 CCTGAGATCGTCATGTGA 1086  
|||||  
Db 1078 CCTGCCATTATCATGTGA 1095

## RESULT 4

US-10-032-585-6223  
; Sequence 6223, Application US/10032585  
; Publication No. US20030180953A1  
; GENERAL INFORMATION:  
; APPLICANT: Terry, Roemer D.  
; APPLICANT: Bo, Jiang  
; APPLICANT: Charles, Boone  
; APPLICANT: Howard, Bussey  
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery  
; FILE REFERENCE: 10182-005-999  
; CURRENT APPLICATION NUMBER: US/10/032,585  
; CURRENT FILING DATE: 2001-12-20  
; NUMBER OF SEQ ID NOS: 8000  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 6223  
; LENGTH: 1089  
; TYPE: DNA  
; ORGANISM: Candida albicans  
US-10-032-585-6223

Query Match 28.6%; Score 310.8; DB 12; Length 1089;  
Best Local Similarity 57.2%; Pred. No. 9.2e-93;  
Matches 624; Conservative C; Mismatches 457; Indels 9; Gaps 3;

QY 1 ATGAAGCCCTCATTTCTGTGGGGTTTTCGGAACCGGCTTCGGCTTTGACTCTGACC 60  
|||||  
Db 1 ATGAAGAGGATTAATTTAGTTCGGAGGATACGGTACCAGATTGAGACCATTAACTTTAACA 60  
  
QY 61 TTCCCGAAACCCCTCGTGAATTTTGAACAACGCCCATGATTCTGCACAGATCGAAGCT 120  
|||||  
Db 61 TTACCAAAACCATTTGGTCGAGTTTCGTACAGACCAATGATCTTGACCAATCGAAGCT 120  
  
QY 121 TTGAAGAAGTTGGGTACAGAGTGGTTTGGCTATCAACTATCGCCAGAGTAATG 180  
|||||  
Db 121 TTGGCCGCTGCTGCTCACCGATATTTGTTTGGCGTCAACTACCGTCCAGAGTCATG 180  
  
QY 181 ATTAATTTCTTGAAGACTTTTGAAGTAAGCTTGGCATCACAAATACATCTCCCAAG 240  
|||||  
Db 181 GTTCCACTTTAAGAAATACGAAGAAGATACGGTGTCTCCATCACCTTCTGTGAA 240  
  
QY 241 ACTGAGCCCTTAGGAACCGCTGCGCCTCTTGTCTTAGCAAGGACAAAGCTTCGGATCGA 300  
|||||  
Db 241 GAAGAACCCTTTGGGCACCGCGCTCCATTAAAGTTGGCTGAAGAAGTGT---GAAAAA 297  
  
QY 301 TCTGGCCAGCCATTTCTTCTCCTACAGTGTGTCATAAGCGAATACCCATTCTCTGAA 360  
|||||  
Db 298 GAGACTCACCATTCTTCTGCTTGAACCTCGAGTCAATTTGGCACTACCCATTCAAGGAA 357  
  
QY 361 CTCATCAAAATTTCAAGTGTCTATGTTGGTGTAGGCAACAATATGTCACTAAGTGGAT 420  
|||||  
Db 358 TTGGCCGACTTCACAAGGCCACCGCGCTGCTGTTACTATTGTTGCACCAAGTCCAC 417  
  
QY 421 GAACCATCAAAATACCGTGTGTGGTTATGGA---GGAGGCACTGGCAGGTTGAAAGG 477  
|||||  
Db 418 GAACCATCAAAATACCGGCTCATTTGTCACGACAGACACTCCAAACTTGTATCGACAGA 477  
  
QY 478 TTGTTGAGAAGCCAAAAATATTTTGGGTAAACAAGTCAATGCTGGGATTTTACTTCTG 537  
|||||  
Db 478 TTGTTGAAAAACGAGTCGAGTTTGTGTAACAGAAATTAACCGCGTTTATACATCTG 537  
  
QY 538 AACCATCTCTCTTGAACCGCATGTAGCTGAGGCCAAATCAATTTGAGAAAGAGTCTTC 597  
|||||  
Db 538 AACCATCTGCTCATCGACTTGAATGAAATGAGACCAACCTCAATCGAAAAAGAACTTTC 597  
  
QY 598 CCTCAAAATTCGAGCTGATCAACAGCTCTATGCAATGGTCTTCCAGGTTTTCAGTGGAT 657  
|||||

Db 598 CCAATCTTGGTGGAAAAACAATTGTACTCTCTCGACTTGGAGGTTTACTTGGATGGAC 657  
QY 658 GTTGGTCAGCCTAGGACTACATTACTTGGCTTGGCTTTTATCTAGACTCGATTAGGAAG 717  
|||||  
Db 658 GTGGGTCACCAAAAGATTTCCTTTCTGAAACCTGTTGTACTGTGACGTGCTATTCGAAA 717  
|||||  
QY 718 AAATCAGCTGCCAAGCTAGCTACTGAGCACATGT---TGTGGCAATGTCTGCTGTCAT 774  
|||||  
Db 718 AAACACCCAGAGAAATTTGTGTAAAGAGAAATACGTCATGCGGTAACTGTGTTGATCGAC 777  
|||||  
QY 775 GAGAGGCCCAAGATTGGAGAGAGTTGTCTGATTGGTCTCGCATTCGCAATTGGACCTGGG 834  
|||||  
Db 778 CCAACTGCCAAGATCCACCCATCTCTCTTAATCGGTCCAAACGTCACCATCGGTCCAAAC 837  
|||||  
QY 835 TGTGTTGTGAGAGACGGGCTGAGGCTTTCCCGCTGCACGTCTCGCGGGCGTGGCTATC 894  
|||||  
Db 838 GTTGTGTCGGTGAAGGTGCTAGAAATCCAAAGATCAAGTGTGTTGGCCCACTCCCAAGTC 897  
|||||  
QY 895 AAGAAGCATGCTTGCAATCTCAACAGCAATTTACGGCTGGCACTCAACTGTTGGTCAATGG 954  
|||||  
Db 898 AAAGACCAACGCTGGGTCAAAATCTACCAATTTGTTGGTTGGAACCTCAGAAATTTGGAAGTGG 957  
|||||  
QY 955 GCACGGATAGAAATATGACTATCTCGGGGAGGATGTTCAATGTTGTGTGAGGTGATC 1014  
|||||  
Db 958 GCTAGAACTGAAGGTGTTTACCGTCTTGGGTGACGACGTGGAAGTGAAGAAAGAAATCTAC 1017  
|||||  
QY 1015 AGCAATPGCGGTGTTGTTCTCCACATAAAGAGATCAAGTCAAGCAATTTCTGAAGCTGAG 1074  
|||||  
Db 1018 GTCAACGGTGCCAAGGCTCTTGGCCACACAAATCGATCTCTCTTCAAGCTCGAAAAAGAGTCT 1077  
|||||  
QY 1075 ATCGTCACTGT 1084  
|||||  
Db 1078 ATCAATTATGT 1087  
|||||

## RESULT 5

US-10-128-714-1009  
; Sequence 1009, Application US/10128714  
; Publication No. US2003011903A1  
; GENERAL INFORMATION:  
; APPLICANT: Jiang, Bo  
; APPLICANT: Hu, Weng  
; APPLICANT: Tishkoff, Daniel  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Eroshkin, Alexey M  
; APPLICANT: Lemieux, Sebastien Y  
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and  
; FILE REFERENCE: 10182-018-999  
; \* CURRENT APPLICATION NUMBER: US/10/128,714  
; CURRENT FILING DATE: 2002-04-23  
; PRIOR APPLICATION NUMBER: US 60/285,697  
; PRIOR FILING DATE: 2001-04-23  
; PRIOR APPLICATION NUMBER: US 60/287,066  
; PRIOR FILING DATE: 2001-04-27  
; PRIOR APPLICATION NUMBER: US 60/295,890  
; PRIOR FILING DATE: 2001-06-05  
; PRIOR APPLICATION NUMBER: US 60/303,899  
; PRIOR FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: US 60/316,362  
; PRIOR FILING DATE: 2001-08-31  
; NUMBER OF SEQ ID NOS: 8603  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1009  
; LENGTH: 1051  
; TYPE: DNA  
; ORGANISM: Aspergillus fumigatus  
US-10-128-714-1009

Query Match 22.8%; Score 247.6; DB 14; Length 1051;  
Best Local Similarity 57.9%; Pred. No. 1.3e-71;  
Matches 521; Conservative 0; Mismatches 364; Indels 15; Gaps 4;

174	QY	174	GGTAAATGATTAAATTTCTTTGAAGGACTTTGAGGATAAGCTTGGCAATCACAATTAATCATGCTC	233
70	DB	70	GCTAAATCTGCAATTTGTGCCCGCAGTAGACGAGGAAACAATACAACTGCAAGTTCGAGTTCTC	129
234	QY	234	CCAAGAGACTGAGCCCTTAGSAAACCGCTGGCCCTCTTGCTCTAGCAAGGGAACAGCTTGC	293
130	DB	130	CGTCCGAGTCCGAACCCCTCGGTACCGCTGTGCTCCCTGAGCTGGCA--GAGAAGATTTT	186
294	QY	294	GGATGGATCTGGCCAGCCATCTTTTGTGCTCAACAGTGAATGTCTATAAGCGAATACCCATT	353
187	DB	187	GGGCAAGAGACGATTTCTCCCTCTTTCTGTTCTCAACTCCGATATCATCTGTGATATCCCTT	246
354	QY	354	TGCTGAACTCATCAAAATTCACAAGTGTCAATGTTGGTGGAGGCAACAATATATGTCACATA	413
247	DB	247	TAAACAGCTCGCAGAGTTCCAAGAAGAACATGGCGACGAGGGCACCATCTGTTGTTACCAA	306
414	QY	414	GGTGGATGAACCATCAAAATACGGTGTGTGGT---TATGAGGAGGCAACTGCGACGGT	470
307	DB	307	GGTCGACGAGCCCTCGAAGTACGGTGTGTTGTCCACAAGCCCAACCATCCCTCTCGCAT	366
471	QY	471	GGAAAGGTTTGTGAGAAGCCAAAATAATTTGTGGGTAAACAAGATCAATGCTGGGATTTA	530
367	DB	367	TGATCTGTTTGTGAGAGGCGAGTTGAGTTGCGTTGGTAACCGCAATTAACGCCGGATATCTA	426
531	QY	531	CTTACTCAACCCCATCTGCTCTGACCAGATTGAGCTGAGGCCAACATCAATTTGAGAAAGA	590
427	DB	427	CATCCTCAACCCCTAGTGTCTCAAGCGCATTTGAGCTGGTCTTACTCCATCGAACAGGA	486
591	QY	591	GGTCTTCCCTCAAAATTCGAGCTGATCAACAGCTCTATGCAATGTGCTCTCCAGGTTTTTG	650
487	DB	487	GACATTTCCCGCCCATCTGCAACGAGCGTCACTCCCACTCTTTGATCTCGAGGGTTTTCTG	546
651	QY	651	GATGGATTTGGTCAGCCCTAAGGACTACATTACTGGCTTGGCTCTTTATCTAGACTCGAT	710
547	DB	547	GATGGATTTGGTCAACCCCAAGATTTCTGACGGGCACTGCTCTTACCTCGCTCGCT	606
711	QY	711	TAGGAAGA-----AATCAGCTGCCAAGCTAGCTACTGGAGCACATGT-TGTTGGCAA	761
607	DB	607	CGCGAAGCGTAATCTCCAAGCTGTGCGCCCAACAGCGAGCCGTACGTTACGCGGCA	666
762	QY	762	TGTGCTGTGCTGATGAGAGCGCAAGATTGGAGAGGTTGTCGTGATTGGTCTGTATGTCGC	821
667	DB	667	CGTCATGGTTGATCCCTCGGCAAGATCGGCAAGAACTGTCCGCAATTTGGCCCTAATGTAGT	726
822	QY	822	CATTGGACCTGGGTGTGTTGGAGGACGGCTGAGGCTTTCCCGCTGCACTGTCAATCG	881
727	DB	727	CATTGGTCCCAACGTTGTGTCGCGATGGCTGGCTCTGCAACGCTGTGTGCTCTTGA	786
882	QY	882	CGGCGTCGGTATAAGAAAGCATGCTTGCATCTCAACAGCAATTATCGCTGGCACTCAAC	941
787	DB	787	GAACAGCAAGGTCGAAGACCAATGCTTGGATCAAGTCGACTATTGTCGGTTGGAACAGCTC	846
942	QY	942	TGTTTGTCAATGGCACGGATAGAGAATATGACTATCTCTGGGAGAGGATGTTTCATGTGTG	1001
847	DB	847	CGTTGGCAAGTGGCTCGCTTGGAGAATGTCAACGGTCTTGGGTGACGATGTACCATTTGC	906
1002	QY	1002	TGATGAGGTGTACAGCAATGCGGTGTGTTCTCCCATAAAGAGATCAAGTCAAGCAT	1061
907	DB	907	TGACGAGGTGTATGTAATGACGGCTCTATTCTCCGCAACGAAGAGCATCAAGACAGACAT	966

## RESULT 6

```

US-10-128-714-6009
; Sequence 6009, Application US/10128714
; Publication NO. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wenqi
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Eroshkin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and

```

QY 882 CGCGTGGTATCAAGAGCATGCTTGCATCTCAACAGCATTATCGCTGGCACTCAAC 941  
 Db 1106 GAACAGCAAGGTCAAGAGCATTGCTTGGATCAAGTCGATATTGTCGGTGGAAAGCTC 1165  
 QY 942 TGTGTGTAATGGGACCGATAGAGTAATGACTATCTCTGGGGAGGATGTTCAATGTG 1001  
 Db 1166 CGTTGCAAGTGGCTCGCTTGGAGAAATGTCACGGTCTTGGGTGACGATGTCAACCATTTGC 1225  
 QY 1002 TGATGAGGTGTACAGCAATCGGGGTGTTGTTCTCCCAATAAAGAGATCAAGTCMAGCAT 1061  
 Db 1226 TGACGAGGTGTATGTCAATCGCGGCTCTATTCTGCCCAAGAGCATCAAGCAGAAACAT 1285  
 RESULT 7  
 US-10-128-714-9  
 ; Sequence 9, Application US/10126714  
 ; Publication No. US20030119013A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Jiang, Bo  
 ; APPLICANT: Hu, Wenqi  
 ; APPLICANT: Tishkoff, Daniel  
 ; APPLICANT: Zamudio, Carlos  
 ; APPLICANT: Eroshkin, Alexey M  
 ; APPLICANT: Lemieux, Sebastien M  
 ; TITLE OF INVENTION: Identification of Essential Genes in *Aspergillus fumigatus* and  
 ; FILE OF INVENTION: Methods of Use  
 ; FILE REFERENCE: 10182-018-999  
 ; CURRENT APPLICATION NUMBER: US/10/128,714  
 ; CURRENT FILING DATE: 2002-04-23  
 ; PRIOR APPLICATION NUMBER: US 60/285,697  
 ; PRIOR FILING DATE: 2001-04-23  
 ; PRIOR APPLICATION NUMBER: US 60/287,066  
 ; PRIOR FILING DATE: 2001-04-27  
 ; PRIOR APPLICATION NUMBER: US 60/295,890  
 ; PRIOR FILING DATE: 2001-06-05  
 ; PRIOR APPLICATION NUMBER: US 60/303,899  
 ; PRIOR FILING DATE: 2001-07-09  
 ; PRIOR APPLICATION NUMBER: US 60/316,362  
 ; PRIOR FILING DATE: 2001-08-31  
 ; NUMBER OF SEQ ID NOS: 8603  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 9  
 ; LENGTH: 3051  
 ; TYPE: DNA  
 ; ORGANISM: *Aspergillus fumigatus*  
 US-10-128-714-9  
 Query Match 22.8%; Score 247.6; DB 14; Length 3051;  
 Best Local Similarity 57.9%; Pred. No. 2.4e-71;  
 Matches 521; Conservative 0; Mismatches 364; Indels 15; Gaps 4;  
 QY 174 GGTAAATGTAATTTCTTGAAGGACTTTGAGATAAGCTTGGCATCAACATTAATGCTC 233  
 Db 1070 GCTAATCTGAATTTGCGCCGAGTACGAGGAACAATACACGTCAGAAATCGATTCTC 1129  
 QY 234 CCAAGAGACTAGCCCTTAGAACCGCTGCGCTCTTCTCTAGCAAGGCAAGCTTGC 293  
 Db 1130 CGTCGAGTCGAACCCCTCGGTACCGCTGCTCCCTGGAAGCTGGCA---GAGAAGATTT 1186  
 QY 294 GGATGATCTGGCCAGCCATCTTTGTCTCAACAGTGTATGTCATAAGCGAATACCCATT 353  
 Db 1187 GGGCAAGGAGCATTTCTCCCTCTTCTGTTCTCAACTCCGATATCATCTGTGA---TATCCCTT 1246  
 QY 354 TGCTGAATCATCAATTTTCAAGTGTCTATGCTGTGAGGCAACAAATATGCTCACTAA 413  
 Db 1247 TAGACAGCTCGAGAGTTTCCAAAGAAACATGCGGACGAGGAGCAACATCGTTGTACCAA 1306  
 QY 414 GGTGGATGAACCAATCAAAATACCGTGTGTGTGT---TATGAGGAGGCAACTGGCAGGT 470  
 Db 1307 GGTGCGAGCCCTCGAAGTACGGTGTGTTGTCACAGAGCCCAACCATCTCTCGCAT 1366  
 QY 471 GGAAGGTTTGTGAGAGGCCAAATAATTTGTGGGTAAACAAGATCAATGCTGGGATTTA 530

Db 1367 TGATCGTTCTGAGAGAGCGAGTTGAGTTGTTGGTAACCGCATTAAACGCGGTATCTA 1426  
 QY 531 CTTACTGAACCCATCTGTCTTGACCGCATTAGCTGAGGCAACATCAATTTGAGAAAGA 590  
 Db 1427 CATCTCTCAACCCCTAGTGTCTCAAGCGCATTAGCTGAGTCTACCTCATCGAACAGGA 1486  
 QY 591 GGTCTTTCCCTCAAAATGCGAGCTGATCAACAGCTCTATCAATGGTCTCTCCAGGTTTTC 650  
 Db 1487 GACATTTCCCGGCATCTGCGGAGCGGTAGCTCACTCCCTTGTATCTCGAGGGTTTCG 1546  
 QY 651 GATGATGTTGTGTCAGCCCTAGGACTACATTAATGCTGCTTGTATCTAGACTCGAT 710  
 Db 1547 GATGGATGTTGTCAACCCCAAGAATTCTGAGCGGCACCTGCTCTACCTCGCTCGCT 1606  
 QY 711 TAGGAGA-----ATCAGCTGCCAAGCTAGCTACTGAGAGCATGT-TGTTGGCAA 761  
 Db 1607 CGCAAGCGTAATCAAGCTCTGCGCCCAACAGCGAGCGGTACGTTACGCGCGCAA 1666  
 QY 762 TGTGCTGTTGCATGAGAGCGCAAGATTGGAGAAAGTGTCTGATTTGGTCTCTGATGTCG 821  
 Db 1667 CGTCATGTTGATCCTCGGCCAAGATCGCAAGAACTGTGCGCATTTGCCCTTAATGATG 1726  
 QY 822 CATTGAGCCTGGTGTGTTGTCAGAGCGGCTGAGGCTTTCCCGCTGACTCTCATCGG 891  
 Db 1727 CATTGCTCCCAAGCTGTGTTGTCGCGGATGGGCTGCTGCAACGCTGTGTGCTTTGGA 1786  
 QY 882 CGCGTGGCTATCAAGAGCATGCTTTGATCTCAAAAGAGCATTTATCGCTGGCACTCAAC 941  
 Db 1787 GAACAGAGGTCAAGGACCATGCTTTGGATCANGTCGACTATTGTCGGTTGMAACAGCTC 1846  
 QY 942 TGTGCTCAATGGGACCGATAGAGAAATGACTATCTCTGGGGAGGATGTTTCAATGTTG 1001  
 Db 1847 CSTTGGCAAGTGGCTCGCTTGGAGAAATGTCACGGTCTTGGGTGACGATGTCAACATTC 1906  
 QY 1002 TGATGAGGTGTACAGCAATGGCGGTGTTGTTCTCCACATAAAGAGATCAAGTCAAGCAT 1061  
 Db 1907 TGACGAGGTGTATGTCATGTCGGGCTCTATTCTGCCCAAGAGCATCAAGCAACAT 1966  
 RESULT 8  
 US-10-128-714-5009  
 ; Sequence 5009, Application US/10128714  
 ; Publication No. US20030119013A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Jiang, Bo  
 ; APPLICANT: Hu, Wenqi  
 ; APPLICANT: Tishkoff, Daniel  
 ; APPLICANT: Zamudio, Carlos  
 ; APPLICANT: Eroshkin, Alexey M  
 ; APPLICANT: Lemieux, Sebastien M  
 ; TITLE OF INVENTION: Identification of Essential Genes in *Aspergillus fumigatus* and  
 ; FILE OF INVENTION: Methods of Use  
 ; FILE REFERENCE: 10182-018-999  
 ; CURRENT APPLICATION NUMBER: US/10/128,714  
 ; CURRENT FILING DATE: 2002-04-23  
 ; PRIOR APPLICATION NUMBER: US 60/285,697  
 ; PRIOR FILING DATE: 2001-04-23  
 ; PRIOR APPLICATION NUMBER: US 60/287,066  
 ; PRIOR FILING DATE: 2001-04-27  
 ; PRIOR APPLICATION NUMBER: US 60/295,890  
 ; PRIOR FILING DATE: 2001-06-05  
 ; PRIOR APPLICATION NUMBER: US 60/303,899  
 ; PRIOR FILING DATE: 2001-07-09  
 ; PRIOR APPLICATION NUMBER: US 60/316,362  
 ; PRIOR FILING DATE: 2001-08-31  
 ; NUMBER OF SEQ ID NOS: 8603  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 5009  
 ; LENGTH: 3370  
 ; TYPE: DNA  
 ; ORGANISM: *Aspergillus fumigatus*  
 US-10-128-714-5009



Query Match 22.8%; Score 247.6; DB 14; Length 3370;  
 Best Local Similarity 57.9%; Pred. No. 2.6e-71;  
 Matches 521; Conservative 0; Mismatches 364; Indels 15; Gaps 4;

Qy 174 GGTAAATGATTAATTTCTTGAAGGACTTTGAGCAATAGCTTGGCATCAAAATACATGCTC 233  
 Db 1389 GCTAATCTGCAATTTGTCGCCGCGATGAGGAGCAATACAGCTGAGATCGATGTTCTC 1448

Qy 234 CCAAGAGACTAGCCCTTAGAACCGTGGCCCTCTTGTCTCTAGCAAGGAGCAAGTTGC 293  
 Db 1449 CGTCAGTCCGAACCCCTCGTACCGTGGTCCCTCAAGCTGCA---GAGAAGATTT 1505

Qy 294 GGATGGATCTGGCCAGCATCTTTGCTCAACAGTGTCTATAGCAATACCCATT 353  
 Db 1506 GGGCAAGACGATTTCTCCCTTCCTGCTCAATCCGATATCATCTGTGATTATCCCTT 1565

Qy 354 TGCTGAATCATCAAAATTTCAAAAGTGTCTAGTGGTGAGGCAACAATTTATGGTCACTAA 413  
 Db 1566 TAAGCAGCTCGCAGATTTCCACAGAAACATGGCGAGGCGACCATCTGTTGTACCA 1625

Qy 414 GGTGGATGAACCATCAAAATACGCTGTTGGT---TATGGAGAGGCAACTGGCAGGGT 470  
 Db 1626 GGTGAGAGCCCTCGAAGTACGCTGCTGTTGTCCACAAGGCCCAACCATCCCTCTCGAT 1685

Qy 471 GGAAAGGTTGTTGAGAGCCAAAATAATTTGTGGTAAACAAGATCAATGCTGGATTTA 530  
 Db 1686 TGATCGTTTGTGTCAGAGCCAGTTGATGTTGTTGTTAAACCGCATTAACCCGGTATCTA 1745

Qy 531 CTACTAGCAACCATCTGTCTTGAACCGCATTGAGTGGTGAGGCAACAATCAATTTGAGAAGA 590  
 Db 1746 CATCTCAACCTAGTGTCTCAAGCGCATTTGAGCTGGTCTTACCTCCATCGNACAGGA 1805

Qy 591 GGTCTTCCCTCAAAATGTCAGTGTGATCAACAGCTCTATGCAATGTCCTTCAGGTTTGTG 650  
 Db 1806 GACATTCCTCCGCCATCTGCAACGAGCGGTGAGCTCCACTCTCTTATCTCGAGGGTTCGTG 1865

Qy 651 GATGGATGTTGGTCAGCTAGGACTACATCTACTGCTGGCTTTTATCTAGACTCGAT 710  
 Db 1866 GATGGATGTTGGTCAACCCAAAGATTTCTGACGGGCACTGCTCTACTCTGCTCGCT 1925

Qy 711 TAGGAAGA-----AATCAGCTGCCAAGCTAGCTACTGAGCACATGT-TGTTGGCAA 761  
 Db 1926 CGCAAGCGTAATCTCAAGCTGCTGGCCCCAAACAGGAGCGGTACTCTACGGCGCAA 1985

Qy 762 TGTGCTGTGATGAGAGCG-CAAGATTTGGAGAGGTTGTCTGATGGTCTGATGTCG 822  
 Db 1986 CGTCAATGTTGATCCCTCGG-CAAGATCGGCAAGAACTGTGCGATTGGCCCTAATGTAGT 2045

Qy 822 CATGGACCTGGGTGTTGTGGAGGAGCGGTGAGGCTTTCCGCTGCACTGTCATGCG 881  
 Db 2046 CATTTGTTCCCAACGTTGTGTCGGGATGGGCTGCTGTGCAAGCTGTGCTCTTGGGA 2105

Qy 882 CGGCTGGTATCAAGAAGCATGCTTGCACTCTCAACAGCATTTATCGGCTGGCACTCAAC 941  
 Db 2106 GAACAGCAAGGTCAAGACCATGCTTGGATCAAGTGCAGCTATTGCTGGTTGGACAGCTC 2165

Qy 942 TGTGGTCAATGGGACGGATAGAGATATGACTATCTCTGGGGAGAGATGTTGATGTTG 1001  
 Db 2166 CGTTGGCAAGTGGGCTGGCTTGGAGATGTCAACGCTCTTGGGTGACGATGTCAACATTGC 2225

Qy 1002 TGATGAGGTACAGCAATGCGGTGTTGTTCTCCACATTAAGAGATCAAGTCAAGCAT 1061  
 Db 2226 TGACAGGTGTATGTCAATGCGGCTCTATTCTGCCCCCAAGAGCATCAAGCAGAACAT 2285

## RESULT 9

US-10-128-714-2009  
 ; Sequence 2009, Application US/10128714  
 ; Publication No. US20030119013A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Jiang, Bo  
 ; APPLICANT: Hu, Wenqi

APPLICANT: Tishkoff, Daniel  
 APPLICANT: Zamudio, Carlos  
 APPLICANT: Eroshkin, Alexey M  
 APPLICANT: Lemieux, Sebastien M  
 TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and  
 FILE REFERENCE: 10182-018-999  
 CURRENT APPLICATION NUMBER: US/10/128,714  
 CURRENT FILING DATE: 2002-04-23  
 PRIOR APPLICATION NUMBER: US 60/285,697  
 PRIOR FILING DATE: 2001-04-23  
 PRIOR APPLICATION NUMBER: US 60/287,066  
 PRIOR FILING DATE: 2001-04-27  
 PRIOR APPLICATION NUMBER: US 60/295,890  
 PRIOR FILING DATE: 2001-06-05  
 PRIOR APPLICATION NUMBER: US 60/303,899  
 PRIOR FILING DATE: 2001-07-09  
 PRIOR APPLICATION NUMBER: US 60/316,362  
 PRIOR FILING DATE: 2001-08-31  
 NUMBER OF SEQ ID NOS: 8603  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO 2009  
 LENGTH: 960  
 TYPE: DNA  
 ORGANISM: Aspergillus fumigatus  
 US-10-128-714-2009

Query Match 22.2%; Score 241.6; DB 14; Length 960;  
 Best Local Similarity 58.5%; Pred. No. 1.2e-69;  
 Matches 500; Conservative 0; Mismatches 339; Indels 15; Gaps 4;

Qy 220 ACAATTCATGCTCCCAAGAGACTGAGCCCTTAGAAACGGTGGCCCTCTTGCTTAGCA 279  
 Db 25 AGAATCGAGTTCTCGTGGAGTCCGAACCCCTCGTACCGTGGTCCCTGAAGCTGCA 84

Qy 280 AGGCAACCTTGGCGATGGATCTGGCCAGCCATTCTTTGTCTCAACAGTGCATCTATA 339  
 Db 85 ---GAGAAGATTTGGCAAGAGCATCTCCCTTCTGTTCTCACTCCGATATCATC 141

Qy 340 AGCGAATACCCATTTGCTGAATCATCAAAATTCACAAAGTGTGATGGTGGAGGCAACA 399  
 Db 142 TGTGATATCCCTTTAAGCAGCTCGCAGAGTTCCACAAGAAACATGGCGACGAGGCACC 201

Qy 400 ATTATGTCACCTAAGTGGATGAACATCAAAATACGCTGTTGTGT---TATGAGGAG 456  
 Db 202 ATCGTGTGTACCAAGTTCGACGAGCCCTCGAAGTACGCTGCTGTCCACAAAGCCCAAC 261

Qy 457 GCAACTGGCAGGTTGGAAGGTTTGTGAGAAAGCCAAAATAATTTGTGGTTAAACAAGATC 516  
 Db 262 CATCCCTCTCGCATTGATCGTTCGTCGAGAGCCAGTTGAGTTCGTTGGTAAACGCAAT 321

Qy 517 AATGTGGGATTTACTACTGAACCATCTGCTTGTACCCGATTTGAGCTGAGGCGCAACA 576  
 Db 322 AACGCGGTATCTACATCTCAACCTAGTGTCTCAAGCGCATTTGAGCTGCTCTACC 381

Qy 577 TCAATGTGABAGAGGTCTTCCCTCAAAATTCAGCTGATCAACAGCTCTATGCAATGTC 636  
 Db 382 TCATCGAACAGGACATTTCCGCCCATCTGTCAGCGCAGCTCAGCTCCATCTCTTTGAT 441

Qy 637 CTTCCAGGTTTGTGATGATGTTGTCAGCTCAGGCACTACATTAATCGCTTGGCTCTT 696  
 Db 442 CTCGAGGTTTCTGGATGATGTTGGTCAACCAAGATTTCTGACGCGGCACTTGCCTC 501

Qy 697 TATCTAGACTCGATTAGGAAGA-----AATCAGCTGCCAAGCTAGCTACTGAGGAC 748  
 Db 502 TACCTCGCTCGTCCGAAAGCGTAACCTCAAGCTGCTGGGCCCAACAGCGAGCCGTAC 561

Qy 743 ATGT-TGTTGGCAATGCTGCTGTCATGAGCGCCAAAGATTGGAGAGGTTGCTGATT 807  
 Db 562 GTCTACGCGCGCAACGTCATGGTTGATCCCTCGGCAAGATCGGCAAGAACTGTGCAAT 621

Qy 808 GGTCTGTGTCGCCATTCGACCTGGGTGTGTTGTGAGGAGCGGCTGAGGCTTTCCCGC 867

Db 622 GCCCCTAATGTAGTCAATTGGTCCCAACGTTGTGGTGGCGATGGCGTGGCTGTGCAACGC 681  
QY 868 TGCACCTGTCAATGCGCGGGTGGTATCAAGAAAGCATGCTTGCATCTCAAAACAGCACTATC 927  
Db 682 TGTGTGCTCTTGGAGAACAGCAGTCAAGGACATGCTTGGATCAAGTGCATTTGTC 741  
QY 928 GCGTGGCACTCAACTGTTGGTCAATGGGCAAGATAGAGAAATAGCTATCTCTGGGGAG 987  
Db 742 GGTGGAAACAGCTCCGTTGGCAAGTGGGCTCGCTTGGAGAAATGTACGGTCTTGGGTGAC 801  
QY 988 GATGTTCAATGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1047  
Db 802 GATGTCACCAATGCTGACGAGTGTATGTCATGCGCGCTCTATTCTGCCGCCCAAGAGC 861  
QY 1048 ATCAAGTCAAGCAT 1061  
Db 862 ATCAAGCAGAACAT 875

RESULT 10  
US-09-294-093B-2164  
; Sequence 2164, Application US/09294093B  
; Patent No. US20010051335A1  
; GENERAL INFORMATION:  
; APPLICANT: Ito, Laura, Y.  
; APPLICANT: Sherman, Bradley, K.  
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN TASSEL  
; FILE REFERENCE: PL-0009 US  
; CURRENT APPLICATION NUMBER: US/09/294,093B  
; CURRENT FILING DATE: 1999-04-16  
; PRIOR APPLICATION NUMBER: 60/082,567  
; PRIOR FILING DATE: April 21, 1998  
; NUMBER OF SEQ ID NOS: 6207  
; SOFTWARE: PERL Program  
; SEQ ID NO 2164  
; LENGTH: 212  
; TYPE: DNA  
; ORGANISM: Zea mays  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. US20010051335A1 700345651H1  
; NAME/KEY: unsure  
; LOCATION: 129  
; OTHER INFORMATION: a, t, c, g, or other  
US-09-294-093B-2164

Query Match 17.2%; Score 187; DB 9; Length 212;  
Best Local Similarity 98.6%; Pred. No. 8,6e-52;  
Matches 209; Conservative 0; Mismatches 1; Indels 2; Gaps 2;  
QY 558 CATTGAGCTGAGGCCAACATCAATTTGAGAAAGAGTCTTCCCT-CAATTTGCACTGATC 616  
Db 2 CATTGAGCTGAGGCCAACATCAATTTGAGAAAGAGTCTTCCCTGCAATTTGCACTGATC 61  
QY 617 AACAGCTCTATGCAATGCTCTCCAGTCTTTTGATGATGATGATGATGATGATGATGATGAT 676  
Db 62 AACAGCTCTATGCAATGCTCTCCAGTCTTTTGATGATGATGATGATGATGATGATGATGAT 121  
QY 677 ACATTACTGGCTTGGCTCTTTTATCTAGACTCGATTAGGAAGAAATCAGCTGCCAAGCTAG 736  
Db 122 ACATTACNGGC-TGGCTCTTTTATCTAGACTCGATTAGGAAGAAATCAGCTGCCAAGCTAG 180  
QY 737 CTACTGGAGCACATGTTGTTGGCAATGTGCTG 768  
Db 181 CTACTGGAGCACATGTTGTTGGCAATGTGCTG 212

RESULT 11  
US-09-770-961-719/c  
; Sequence 719, Application US/09770961  
; Publication No. US20030115639A1  
; GENERAL INFORMATION:

; APPLICANT: Gorlach, Jorn  
; APPLICANT: An, Yong-Qiang  
; APPLICANT: Hamilton, Carol M.  
; APPLICANT: Price, Jennifer L.  
; APPLICANT: Raines, Tracy M.  
; APPLICANT: Yu, Yang  
; APPLICANT: Rameaka, Joshua G.  
; APPLICANT: Page, Amy  
; APPLICANT: Matthew, Abraham V.  
; APPLICANT: Ledford, Broeze J.  
; APPLICANT: Woessner, Jeffrey P.  
; APPLICANT: Haas, William David  
; APPLICANT: Garcia, Carlos A.  
; APPLICANT: Kricker, Maja  
; APPLICANT: Slader, Ted  
; APPLICANT: Davis, Keith R.  
; APPLICANT: Allen, Keith  
; APPLICANT: Hoffman, Neil  
; APPLICANT: Hurlan, Patrick  
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis  
; FILE REFERENCE: 2026 (PARA-015PRV)  
; CURRENT APPLICATION NUMBER: US/09/770,961  
; CURRENT FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: 60/178,466  
; PRIOR FILING DATE: 2000-01-27  
; NUMBER OF SEQ ID NOS: 999  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 719  
; LENGTH: 488  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-770-961-719

Query Match 16.6%; Score 180.8; DB 11; Length 488;  
Best Local Similarity 80.3%; Pred. No. 1.7e-49;  
Matches 212; Conservative 0; Mismatches 52; Indels 0; Gaps 0;  
QY 823 ATTGACCTGGGTGTGTGTGGAGGACGGGTGAGGCTTTCCCGCTGCACGTGTCAATGCGC 882  
Db 488 ATTGTCCAGGCTGCATTTGTGAGTCAGAGTCAGACCTCTCCGATGCACGTCATGGT 429  
QY 883 GCGTGGCTATCAAGAACGTCCTGCATCTCAACAGCATTCGGCTGGCACTCAACT 942  
Db 428 GGAGTCGGCATCAAGAACGTCGGTGTATCTCGAGCAGTATCATCGGTGGCACTCAACG 369  
QY 943 GTTGTCAATGGGCACGATAGAGAAATATGATCTCTCGGGGAGGATGTTTCATGTGT 1002  
Db 368 GTTGTCAATGGGCACGATAGAGAAATATGATCTCTCGGTGAGGATGTTTCATGTGAGC 309  
QY 1003 GATGAGGTGATCAGCAATGGCGGTGTTCTTCCACATAAAGAGATCAAGTCAAGCAAT 1062  
Db 308 GATGAGTCTATAGCAATGGAGGAGTGTGTTTGGCCACAAAGGAGATCAAAATCAAAATC 249  
QY 1063 CTGAAGCTTGAGATCGTCATGTGA 1086  
Db 248 TTGAAGCCAGAGATAGTATGTGA 225

RESULT 12  
US-09-878-574-2725  
; Sequence 2725, Application US/09878574  
; Patent No. US20020110548A1  
; GENERAL INFORMATION:  
; APPLICANT: Byrum, Joseph R.  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Thompson, Michael D.  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(15401)B  
; CURRENT APPLICATION NUMBER: US/09/878,574  
; CURRENT FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 09/333,535

; PRIOR FILING DATE: 1999-06-14  
; NUMBER OF SEQ ID NOS: 15775  
; SEQ ID NO 2725  
; LENGTH: 364  
; TYPE: DNA  
; ORGANISM: Glycine max  
; OTHER INFORMATION: Clone ID: LIB3028-021-Q1-B1-C7  
US-09-878-574-2725

Query Match 13.9%; Score 151.4; DB 10; Length 364;  
Best Local Similarity 74.3%; Pred. No. 9.8e-40;  
Matches 191; Conservative 0; Mismatches 66; Indels 0; Gaps 0;  
Qy 1 ATGAAGCCCTCATCTCTGTGGGGGTTTCGGAACCCGCTTCGGCCCTTTGACTCTGAGC 60  
Db 108 ATGAAGCGATGATTCGTGTGGGGGATTTGGAAACAAGGCTGAGGCGACCTGACACTCAGT 167  
Qy 61 TTCCCGAAACCCCTCGTGGATTTTGCAAAACAGCCCATGATTTGCAACAGATCGAAGCT 120  
Db 168 TTCCCTAAGCCTCTGTTGATTTTGCTTAACAAGCCTATGATTTGCAATCAGATAGAAGCC 227  
Qy 121 TTGAAGAAGTTGGGCTCACAGAGTGGTTTGGCTATCAACTATCCGCCAGAGGTAATG 180  
Db 228 CTTAAGGCCATTGGAGTCACTAGGTAGTGTAGCCATCAATTACCAACCAAGAGAGTATG 287  
Qy 181 ATTAATTTCTTGAAGGACTTTGAGGATAAGCTTGGCATCAATTCATCTCCCAAGAG 240  
Db 288 TTGAATTTCTTGAAGGATTTTGAATCAAGCTCGGCATCAAGATCACATGTTCTCAGGAA 347  
Qy 241 ACTGAGCCCTTAGGAAC 257  
Db 348 ACTGAACCACTGGGAAC 364

RESULT 13  
US-09-878-574-8673  
; Sequence 8673, Application US/09378574  
; Patent No. US20020110548A1  
; GENERAL INFORMATION:  
; APPLICANT: Byrum, Joseph R.  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Thompson, Michael D.  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(15401)B  
; CURRENT APPLICATION NUMBER: US/09/878,574  
; PRIOR FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 09/333,535  
; PRIOR FILING DATE: 1999-06-14  
; NUMBER OF SEQ ID NOS: 15775  
; SEQ ID NO 8673  
; LENGTH: 269  
; TYPE: DNA  
; ORGANISM: Glycine max  
; OTHER INFORMATION: Clone ID: 701101510H1  
US-09-878-574-8673

Query Match 12.0%; Score 130.6; DB 10; Length 269;  
Best Local Similarity 70.3%; Pred. No. 7.4e-33;  
Matches 189; Conservative 0; Mismatches 79; Indels 1; Gaps 1;  
Qy 230 GCTCCCAAGAGACTGAGCCCTTAGGAACCGCTGGCCCTCTTGCTCTAGCAAGGCAAGC 289  
Db 1 GTTCTCAGGAATCAAGCACTGGGAACAGCAGGTCCTCTGCTCTTGGCCAGGATAGC 60  
Qy 230 TTGGGATGGATTCGGCC-AGCCATCTTTTGTCTCAACAGTATGTCATAAGCGAATAC 348  
Db 61 TGATAGATGACTCTGGAGAACCTTTTGTCTCAACAGTATGTTATCAGTGAGTAT 120  
Qy 349 CCATTGCTGAACTCATCAATTTTCAACAGTGTCTGCTGGTGAAGCAAAATATCGTC 408  
Db 121 CCATCAAGAAGTAATGATTGATTTCCATTAACAAACCATGGAGGAGGCTTCCATAATGGTA 180

Qy 409 ACTAAGTGGATGAACCATCAAAATACGGTGTGTGGTTATGGAGGAGGCAACTGGCAGG 468  
Db 181 ACCAAGTTTGACGAGCCATCAAAAGTACGGCTGGTTGTGATGGGAAGAGACACAGGCGAG 240  
Qy 469 GTGGAAAGGTTTGTGAGAAGCCAAAT 497  
Db 241 GTTGATAAATTTGTGAAAAACCGAATTT 269

RESULT 14  
US-09-878-574-10462  
; Sequence 10462, Application US/09878574  
; Patent No. US20020110548A1  
; GENERAL INFORMATION:  
; APPLICANT: Byrum, Joseph R.  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Thompson, Michael D.  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(15401)B  
; CURRENT APPLICATION NUMBER: US/09/878,574  
; PRIOR FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 09/333,535  
; PRIOR FILING DATE: 1999-06-14  
; NUMBER OF SEQ ID NOS: 15775  
; SEQ ID NO 10462  
; LENGTH: 272  
; TYPE: DNA  
; ORGANISM: Glycine max  
; OTHER INFORMATION: Clone ID: 700966702H1  
US-09-878-574-10462

Query Match 7.3%; Score 79.2; DB 10; Length 272;  
Best Local Similarity 72.9%; Pred. No. 1.2e-15;  
Matches 102; Conservative 0; Mismatches 38; Indels 0; Gaps 0;  
Qy 1 ATGAAGGCCCTCATCTCTGTGGGGGTTTCGGAACCCGCTTCGGCCCTTTGACTCTGAGC 60  
Db 132 ATGAAGGCATTCATCTCTGTGGGGGTTTGGAAACAGGCTGAGGCCACTGACACTCAGT 191  
Qy 61 TTCCCGAAACCCCTCGTGGATTTTGCAAAACAGGCCATGATTTGCAACAGATCGAAGCT 120  
Db 192 TTCTCTAAGCCCTCTGTTGATTTGCTAACAACCTAGATTCTGCATCAGATAGAGGCC 251  
Qy 121 TCGAAGAAGTTGGGCTCAC 240  
Db 252 CTTAAGGCCATTTGGAGTTAC 271

RESULT 15  
US-10-156-761-6955  
; Sequence 6955, Application US/10.56761  
; Publication No. US20030119018A1  
; GENERAL INFORMATION:  
; APPLICANT: OMURA, SATOSHI  
; APPLICANT: IKEDA, HARUO  
; APPLICANT: ISHIKAWA, JUN  
; APPLICANT: HORIKAWA, HIROSHI  
; APPLICANT: SHIBA, TADAYOSHI  
; APPLICANT: SAKAKI, YOSHIYUKI  
; APPLICANT: HATTORI, MASAHIRO  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-262  
; CURRENT APPLICATION NUMBER: US/10/156,761  
; PRIOR FILING DATE: 2002-05-29  
; PRIOR APPLICATION NUMBER: JP 2001-204089  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: JP 2001-272697  
; PRIOR FILING DATE: 2001-08-02  
; NUMBER OF SEQ ID NOS: 15109  
; SEQ ID NO 6955  
; LENGTH: 2493  
; TYPE: DNA

; ORGANISM: Streptomyces avermectilis  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(2493)  
US-10-156-761-6955

```
Query Match      6.9%; Score 74.6; DB 14; Length 2493;
Best Local Similarity 51.5%; Pred. No. 1.6e-13;
Matches 254; Conservative 0; Mismatches 224; Indels 15; Gaps 3;

Qy 1 ATGAAGGCCCTCATTTGTGGGGTTTCGGAAACCGCCCTTCGGCCCTTTGACTCTGAGC 60
   |||||
Db 1 ATGAAGGCCGTCGTGATGGCGGAGCGGACACAGCCCTTCGCCGATGACTCGAGT 60

Qy 61 TTCCCGAAACCCCTCGTGGATTTCGAAACAAAGCCCATGATTCGACACAGATCGAAGCT 120
   |||||
Db 61 ATGCCCAAGCGCTCTCTGCCCTGTTAAACCGGCCGATCATGGAGCATGTCTCGGCTTA 120

Qy 121 TTGAAGAAGTTGGGGTCACAGAGGTGTTTGGCTATCAACTATCGCCAGAGGTAATG 180
   |||||
Db 121 CTCAAAAGGATGGGCTCAAAGAGACGTCGTAACCTGTGCAGTTCCTGGCGTCTCTGTC 180

Qy 181 ATTAATTTCTTGAAGGACTTTGAGGATAAGCTTGGCATCAATTAATGCTCCCAAGAG 240
   |||||
Db 181 AAGAACTACTTCGGTGAC---GGCGAAGAGCTGGCATGGAGCTCACTTATGCCAATGAG 237

Qy 241 ACTGAGCCCTTAGAAGCGCTGGCCCTCTTGCTCTAGCAAGGACAGCTTGGGATGGA 300
   |||||
Db 238 GAGAAGCCACTCGGTACCGCTGGAGCGTCAAGAACCGCGGAGGAGCGCTGAAGGA---- 293

Qy 301 TCTGGCCAGCCATTCTTTGTCTCAACAGTGATGTCAATAGCGAATACCCATTGTGTAA 360
   |||||
Db 294 -----CGACACTTTCCTCGTCACTCGGGTGAGCCCTGACCGACTTCGATCTCACCGAG 348

Qy 361 CTCATCAAAATTCAAAGTGTCATGGTGGTGAGGCAACAATATGGTCACTAAGGTGGAT 420
   |||||
Db 349 CTGATCAATTTCCACAAGGAAAGGGTGCGCTGTCACCGTCTGTCTGACGCGGTGCC 408

Qy 421 GAACCATCAAAATACGGTGTGTGGTTATGGAGGAGGCAACTGGCAGGGTGGAAAGGTTT 480
   |||||
Db 409 AATCCTCTGGAATTCGGCATACCATCGTCGACGAGAA---GGCAAGGTCGAGCGCTTC 465

Qy 481 GTTGAGAAGCCAA 493
   |||||
Db 466 CTCGAGAAGCCCA 478
```

Search completed: October 15, 2003, 11:00:17  
Job time : 376 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 9, 2003, 14:03:06 ; Search time 31 Seconds  
(without alignments)  
1876.366 Million cell updates/sec

Title: US-09-374-967-2  
Perfect score: 1864  
Sequence: 1 MKALLVGGFGRRLRPLTUS.....GVVLPHEIKSSILKPEIVM 361

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext: 0.5

Searched: 600653 seqs, 161128416 residues

Total number of hits satisfying chosen parameters: 600653

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	IF	Description
1	1681	90.2	361	12	US-10-342-224-24	Sequence 24, Appl
2	1681	90.2	361	12	US-10-427-631-32	Sequence 32, Appl
3	1183.5	63.5	360	12	US-10-427-631-2	Sequence 2, Appl1
4	1161.5	62.3	364	15	US-10-128-714-8009	Sequence 8009, Ap
5	1099.5	59.0	362	12	US-10-032-585-7223	Sequence 7223, Ap
6	1045	56.1	373	12	US-10-427-631-33	Sequence 33, Appl
7	878.5	47.1	319	15	US-10-128-714-3009	Sequence 3009, Ap
8	515	27.6	831	15	US-10-156-761-14505	Sequence 14505, A
9	488	26.2	359	10	US-09-712-363-269	Sequence 269, App
10	469.5	25.2	360	15	US-10-156-761-12571	Sequence 12571, A
11	431.5	23.1	348	10	US-09-738-626-4325	Sequence 4325, Ap
12	314	16.8	355	12	US-09-769-734-33	Sequence 33, Appl
13	296	15.9	355	10	US-09-922-683-10	Sequence 10, Appl
14	281.5	15.1	237	15	US-10-156-761-8552	Sequence 8552, Ap
15	277.5	14.9	453	10	US-09-925-637-18	Sequence 18, Appl

16	277.5	14.9	453	15	US-10-084-205-18	Sequence 18, Appl
17	275.5	14.8	449	9	US-09-815-242-5479	Sequence 5479, Ap
18	275.5	14.8	452	9	US-09-815-242-12149	Sequence 12149, A
19	257.5	13.8	243	15	US-10-156-761-11498	Sequence 11498, A
20	247.5	13.3	237	15	US-10-156-761-7906	Sequence 7906, Ap
21	239	12.8	146	10	US-09-738-973-190	Sequence 190, App
22	239	12.8	146	10	US-09-854-133-190	Sequence 190, App
23	239	12.8	146	15	US-10-144-649A-190	Sequence 190, App
24	228.5	12.3	224	11	US-09-975-719-175	Sequence 175, App
25	227	12.2	456	9	US-09-815-242-10400	Sequence 10400, A
26	216.5	11.6	458	9	US-09-815-242-11759	Sequence 11759, A
27	216	11.6	456	9	US-09-815-242-14064	Sequence 14064, A
28	209	11.2	456	9	US-09-815-242-11073	Sequence 11073, A
29	198.5	10.6	299	15	US-10-156-761-8485	Sequence 8485, Ap
30	195.5	10.4	748	12	US-10-032-585-7314	Sequence 7314, Ap
31	194.5	10.4	285	10	US-09-738-626-3871	Sequence 3871, Ap
32	190	10.2	292	10	US-09-861-289-12	Sequence 12, Appl
33	190	10.2	292	10	US-09-860-846-12	Sequence 12, Appl
34	190	10.2	292	11	US-09-988-184B-12	Sequence 12, Appl
35	190	10.2	292	11	US-09-836-821-12	Sequence 12, Appl
36	190	10.2	292	11	US-09-793-708-14	Sequence 14, Appl
37	190	10.2	292	12	US-10-201-365-16	Sequence 16, Appl
38	190	10.2	292	12	US-10-160-535-14	Sequence 14, Appl
39	190	10.2	3782	10	US-09-861-289-4	Sequence 4, Appl1
40	190	10.2	3782	10	US-09-860-846-4	Sequence 4, Appl1
41	190	10.2	3782	11	US-09-988-184B-4	Sequence 4, Appl1
42	190	10.2	3782	11	US-09-836-821-4	Sequence 4, Appl1
43	189.5	10.2	461	9	US-09-815-242-10937	Sequence 10937, A
44	188	10.1	485	10	US-09-738-626-4551	Sequence 4551, Ap
45	187.5	10.1	456	12	US-10-326-185-30	Sequence 30, Appl

ALIGNMENTS

RESULT 1  
US-10-342-224-24  
; Sequence 24, Application US/10342224  
; Publication No. US20030162294N1  
; GENERAL INFORMATION:  
; APPLICANT: Nathalie Verbruggen  
; TITLE OF INVENTION: Genes Involved in Tolerance to Environmental Stress  
; FILE REFERENCE: CNN-012US  
; CURRENT APPLICATION NUMBER: US/10/342,224  
; CURRENT FILING DATE: 2003-01-13  
; PRIOR APPLICATION NUMBER: US/09/762,154  
; PRIOR FILING DATE: 2002-02-02  
; PRIOR APPLICATION NUMBER: EP 98202634.6  
; PRIOR FILING DATE: 1998-08-04  
; NUMBER OF SEQ ID NOS: 123  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 24  
; LENGTH: 361  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
US-10-342-224-24

Query Match	90.2%	Score 1681	DB 12	Length 361
Best Local Similarity	98.6%	Pred. No. 6e-169		
Matches 320	Conservative 22	Mismatches 19	Indels 0	Gaps 0
QY	1	MKALLVGGFGRRLRPLTUSFPKPLVDFANKPMLHQIEALKEVGVTEVVLAINRPEVM	60	
DB	1	MKALLVGGFGRRLRPLTUSFPKPLVDFANKPMLHQIEALKEVGVTEVVLAINRPEVM	60	
QY	61	INFLKDFEDKLGITITSCQETEPGLTAGPLARLADKLGSGQFFVLSNDVISEYFAE	120	
DB	61	LNFLKDFETKLEIKITSCQETEPGLTAGPLARLADKLGSGQFFVLSNDVISEYPLKE	120	
QY	121	LKPKHCHGGGATINVTKVDEPSKYGVVWMBEATGRVERFVEKPKIFVGNKINAGIYLLN	180	
DB	121	MLEPKHSGGGEASINVTKVDEPSKYGVVWMBEATGRVERFVEKPKLYVGNKINAGIYLLN	180	

```
QY 181 PSVLDRLELRPTSTKEVFPQIAADQOLYAMVLPGFMMVGGQPRDYITGLRLYLDLSIRKK 240
Db 181 PSVLDRLELRPTSTKEVFPQIAADQOLYAMVLPGFMMVGGQPRDYITGLRLYLDLSIRKK 240
QY 241 SAAKLATGAHVGNVNLVHESAKIGEGCLIGPDVAIGPGCVVDEGVRLSRCTVMRGVRIKK 300
Db 241 SAAKLATGAHVGNVNLVHESAKIGEGCLIGPDVAIGPGCVVDEGVRLSRCTVMRGVRIKK 300
QY 301 HACISNSIIIGHSTVGQWARIENMTILGEDVHVCDEVYNSGGVVLPHKKEIKSSILKPEIV 360
Db 301 HACISNSIIIGHSTVGQWARIENMTILGEDVHVCDEVYNSGGVVLPHKKEIKSSILKPEIV 360
QY 361 M 361
Db 361 M 361

RESULT 2
US-10-427-631-32
; Sequence 32, Application US/10427631
; Publication No. US20030175923A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE CORPORATION; TANG, Y. Tom;
; APPLICANT: CORLEY, Neil C.; GUEGLER, Karl J.;
; APPLICANT: BAUGHN, Mariah R.; LAL, Preeti G.;
; APPLICANT: YUE, Henry; HILLMAN, Jennifer L.;
; APPLICANT: AZIMZAI, Yalda
; TITLE OF INVENTION: HUMAN TRANSFERASE PROTEINS
; FILE REFERENCE: PF-0592-1 DIV
; CURRENT APPLICATION NUMBER: US/10/427,631
; CURRENT FILING DATE: 2003-04-23
; PRIOR APPLICATION NUMBER: US 09/786,240
; PRIOR FILING DATE: 2002-03-12
; PRIOR FILING DATE: 1999-09-09
; PRIOR APPLICATION NUMBER: PCT/US99/20989
; PRIOR FILING DATE: 1998-09-10
; PRIOR FILING DATE: 1998-11-04
; PRIOR FILING DATE: 1999-05-11
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PERL Program
; SEQ ID NO 32
; LENGTH: 361
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: GenBank ID No. US20030175923A1 g2642159
US-10-427-631-32

Query Match 90.2%; Score 1681; DB 12; Length 361;
Best Local Similarity 88.68; Pred. No. 6e-169;
Matches 320; Conservative 22; Mismatches 19; Indels 0; Gaps 0;

QY 1 MKALILVGGFGTGLRLPRLTSLSPKPLVDFANKPMILHQIEALKEVGVTEVVLAINYPPEVM 60
Db 1 MKALILVGGFGTGLRLPRLTSLSPKPLVDFANKPMILHQIEALKEVGVTEVVLAINYPPEVM 60
QY 61 INFLKDFEDKLGITITCSQTEPLECTAGLALARDKLADGSGQPPFFVLNSDVISSEYPPAE 120
Db 61 INFLKDFEDKLGITITCSQTEPLECTAGLALARDKLADGSGQPPFFVLNSDVISSEYPPAE 120
QY 121 LKFKHCHGGEATIMVTKVDSPSKYGVVWMEATGRVERFVEKPKIFVGNKINAGIYLLN 180
Db 121 LKFKHCHGGEATIMVTKVDSPSKYGVVWMEATGRVERFVEKPKIFVGNKINAGIYLLN 180
QY 181 PSVLDRLELRPTSTKEVFPQIAADQOLYAMVLPGFMMVGGQPRDYITGLRLYLDLSIRKK 240
Db 181 PSVLDRLELRPTSTKEVFPQIAADQOLYAMVLPGFMMVGGQPRDYITGLRLYLDLSIRKK 240
QY 241 SAAKLATGAHVGNVNLVHESAKIGEGCLIGPDVAIGPGCVVDEGVRLSRCTVMRGVRIKK 300
Db 241 SAAKLATGAHVGNVNLVHESAKIGEGCLIGPDVAIGPGCVVDEGVRLSRCTVMRGVRIKK 300
QY 301 HACISNSIIIGHSTVGQWARIENMTILGEDVHVCDEVYNSGGVVLPHKKEIKSSILKPEIV 360
Db 301 HACISNSIIIGHSTVGQWARIENMTILGEDVHVCDEVYNSGGVVLPHKKEIKSSILKPEIV 360
QY 361 M 361
Db 361 M 361
```

```
Db 241 SPAKLTSGPHVGNVNLVDETAIGEGCLIGPDVAIGPGCVVDEGVRLSRCTVMRGVRIKK 300
QY 301 HACISNSIIIGHSTVGQWARIENMTILGEDVHVCDEVYNSGGVVLPHKKEIKSSILKPEIV 360
Db 301 HACISNSIIIGHSTVGQWARIENMTILGEDVHVCDEVYNSGGVVLPHKKEIKSSILKPEIV 360
QY 361 M 361
Db 361 M 361

RESULT 3
US-10-427-631-2
; Sequence 2, Application US/10427631
; Publication No. US20030175923A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE CORPORATION; TANG, Y. Tom;
; APPLICANT: CORLEY, Neil C.; GUEGLER, Karl J.;
; APPLICANT: BAUGHN, Mariah R.; LAL, Preeti G.;
; APPLICANT: YUE, Henry; HILLMAN, Jennifer L.;
; APPLICANT: AZIMZAI, Yalda
; TITLE OF INVENTION: HUMAN TRANSFERASE PROTEINS
; FILE REFERENCE: PF-0592-1 DIV
; CURRENT APPLICATION NUMBER: US/10/427,631
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: US 09/786,240
; PRIOR FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: PCT/US99/20989
; PRIOR FILING DATE: 1999-09-09
; PRIOR APPLICATION NUMBER: US 60/172,220
; PRIOR FILING DATE: 1999-09-10
; PRIOR FILING DATE: 1998-11-04
; PRIOR FILING DATE: 1998-11-04
; PRIOR APPLICATION NUMBER: US 60/133,642
; PRIOR FILING DATE: 1999-05-11
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PERL Program
; SEQ ID NO 2
; LENGTH: 360
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030175923A1 2682663CD1
US-10-427-631-2

Query Match 63.5%; Score 1183.5; DB 12; Length 360;
Best Local Similarity 61.2%; Pred. No. 2.4e-216;
Matches 221; Conservative 65; Mismatches 74; Indels 1; Gaps 1;

QY 1 MKALILVGGFGTGLRLPRLTSLSPKPLVDFANKPMILHQIEALKEVGVTEVVLAINYPPEVM 60
Db 1 MKALILVGGFGTGLRLPRLTSLSPKPLVDFANKPMILHQIEALKEVGVTEVVLAINYPPEVM 60
QY 61 INFLKDFEDKLGITITCSQTEPLECTAGLALARDKLADGSGQPPFFVLNSDVISSEYPPAE 120
Db 61 EKEMKAQEQRLGIRISMSSHESEPLCTAGLALARDLSE-TADPPFVLNSDVIDCFPPQA 119
QY 121 LKFKHCHGGEATIMVTKVDSPSKYGVVWMEATGRVERFVEKPKIFVGNKINAGIYLLN 180
Db 120 MYQFRRHGHGQESILVTKVEEPSKYGVVWMEATGRVIRHRFVEKPOVFNKINAGIYLLS 179
QY 181 PSVLDRLELRPTSTKEVFPQIAADQOLYAMVLPGFMMVGGQPRDYITGLRLYLDLSIRKK 240
Db 180 PAVLARIQLQPTSTKEVFPPIVAKESQLYAMELQGFMMDIGQKDFLGMCLFLQLSLRQK 239
QY 241 SAAKLATGAHVGNVNLVHESAKIGEGCLIGPDVAIGPGCVVDEGVRLSRCTVMRGVRIKK 300
Db 240 QPERLCSGPGVGNVNLVDPSPARIGQNCISGPNVSLGPGVWEDGVCTRRCTVLRDARIS 299
QY 301 HACISNSIIIGHSTVGQWARIENMTILGEDVHVCDEVYNSGGVVLPHKKEIKSSILKPEIV 360
Db 301 HACISNSIIIGHSTVGQWARIENMTILGEDVHVCDEVYNSGGVVLPHKKEIKSSILKPEIV 360
```

Db 300 HSWLESCIVGRCVQGWVRNMTVLGEDVINDELYLNGASVLPKHSIGSVPEPRII 359

QY 361 M 361

Db 360 M 360

## RESULT 4

US-10-128-714-8009  
; Sequence 8009, Application US/10128714  
; Publication No. US20030119013A1  
; GENERAL INFORMATION:  
; APPLICANT: Jiang, Bo  
; APPLICANT: Hu, Weng  
; APPLICANT: Tishkoff, Daniel  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Eroskin, Alexey M  
; APPLICANT: Lemieux, Sebastien M  
; TITLE OF INVENTION: Identification of Essential Genes in *Aspergillus fumigatus* and  
; FILE REFERENCE: 10182-018-999  
; CURRENT APPLICATION NUMBER: US/-0/128,714  
; CURRENT FILING DATE: 2002-04-23  
; PRIOR APPLICATION NUMBER: US 60/285,697  
; PRIOR FILING DATE: 2001-04-23  
; PRIOR APPLICATION NUMBER: US 60/287,066  
; PRIOR FILING DATE: 2001-04-27  
; PRIOR APPLICATION NUMBER: US 60/295,890  
; PRIOR FILING DATE: 2001-06-05  
; PRIOR APPLICATION NUMBER: US 60/303,899  
; PRIOR FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: US 60/316,362  
; PRIOR FILING DATE: 2001-08-31  
; NUMBER OF SEQ ID NOS: 8603  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 8009  
; LENGTH: 364  
; TYPE: PRT  
; ORGANISM: *Aspergillus fumigatus*  
US-10-128-714-8009

Query Match 62.3%; Score 1161.5; DB 15; Length 364;

Best Local Similarity 60.8%; Pred. No. 5.2e-114;  
Matches 222; Conservative 60; Mismatches 78; Indels 5; Gaps 3;

QY 1 MKALILVGGFGTGLRPLTLSPFKPLVDFANKPMLHQIEALKEVGTVTEVVLAINYRPEVM 60

Db 1 MKALILVGGFGTGLRPLTLSPFKPLVDFANKPMLHQIEALKEVGTVTEVVLAINYRPEVM 60

QY 61 INFLKDFEDKLGITITCSQETPLGTAGPLALARDKLDGSGQPPFFVLNSDVISSEYPPAE 120

Db 61 VAALKKYEQVNVRIEFSVESEPLGTAGPLKLA-EKILGKDDSPFFVLNSDIIICDYPPKQ 119

QY 121 LIKFKHCHGGEATIMVTKVDEPSKYGVVME-EATGVERFVEKPKIFVGNKINAGIYLL 179

Db 120 LAEFHKGHGDGTIVVTKVDEPSKYGVVME-EATGVERFVEKPKIFVGNKINAGIYLL 179

QY 180 NPSVLDRLELRPTSIKEVFPQIAADQOLYAMVLPFGFMDVQGPPOVITGLRLYLDISRK 239

Db 180 NPSVLDRLELRPTSIKEVFPQIAADQOLYAMVLPFGFMDVQGPPOVITGLRLYLDISRK 239

QY 240 KSAAKLATGAHV---GNVLVHESAKIGEGCLIGPDVAIGPGCVVGGVRLSRCTVNRGV 296

Db 240 RNSKULAPNSEPYVYGGNVMDPSAKIGKCRIGPNVVGPNVVGVLQRCVLENS 299

QY 297 RIKKHACISNSIIGHWSTVGCWARTENMTILGEDVHVCDVYNSGGVVLPHKEIKSSILK 356

Db 300 KYKDHAKWIKSTIVGNSSVGVWLRLENVTLGGDDVTIADVYVNGGSIIPKHSIKQNDV 359

QY 357 PEIVM 361

Db 360 PAIIM 364

## RESULT 5

US-10-032-585-7223  
; Sequence 7223, Application US/10032585  
; Publication No. US20030180953A1  
; GENERAL INFORMATION:  
; APPLICANT: Terry, Roemer D.  
; APPLICANT: Bo, Jiang  
; APPLICANT: Charles, Boone  
; APPLICANT: Howard, Bussey  
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery  
; FILE REFERENCE: 10182-005-999  
; CURRENT APPLICATION NUMBER: US/10/032,585  
; CURRENT FILING DATE: 2001-12-20  
; NUMBER OF SEQ ID NOS: 8000  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 7223  
; LENGTH: 362  
; TYPE: PRT  
; ORGANISM: *Candida albicans*  
US-10-032-585-7223

Query Match 59.0%; Score 1099.5; DB 12; Length 362;

Best Local Similarity 57.9%; Pred. No. 1.1e-107;  
Matches 210; Conservative 60; Mismatches 90; Indels 3; Gaps 3;

QY 1 MKALILVGGFGTGLRPLTLSPFKPLVDFANKPMLHQIEALKEVGTVTEVVLAINYRPEVM 60

Db 1 MKGLILVGGYGTGLRPLTLSPFKPLVDFANKPMLHQIEALKEVGTVTEVVLAINYRPEVM 60

QY 61 INFLKDFEDKLGITITCSQETPLGTAGPLALARDKLDGSGQPPFFVLNSDVISSEYPPAE 120

Db 61 VSTLKYEEYGVISITFSEVEEPJGTASPLKAEVLKK-DQSPFFVLNSDVIDCDYPPKE 119

QY 121 LIKFKHCHGGEATIMVTKVDEPSKYGVVME-EAT-GRVERFVEKPKIFVGNKINAGIYLL 179

Db 120 LADPHKAGAAAGTIVATKVDEPSKYGVVME-EAT-GRVERFVEKPKIFVGNKINAGIYLL 179

QY 180 NPSVLDRLELRPTSIKEVFPQIAADQOLYAMVLPFGFMDVQGPPOVITGLRLYLDISRK 239

Db 180 NPSVLDRLELRPTSIKEVFPQIAADQOLYAMVLPFGFMDVQGPPOVITGLRLYLDISRK 239

QY 240 KSAAKLATGAHV-VGNVLVHESAKIGEGCLIGPDVAIGPGCVVGGVRLSRCTVNRGVRI 298

Db 240 XHPEKLCREKYVHGVNVLIDPTAKIHPSALIGENVITGPNVVGEGARIQSVLLANSQV 299

QY 299 KKHACISNSIIGHWSTVGCWARTENMTILGEDVHVCDVYNSGGVVLPHKEIKSSILKPE 358

Db 300 KDHAWKSTIVGNSSRIGKWARTEGTVLGGDDVEVKEIYVNGAKVLPKHSISSNVEKES 359

QY 359 IVM 361

Db 360 IVM 362

## RESULT 6

US-10-427-631-33  
; Sequence 33, Application US/10427631  
; Publication No. US20030175923A1  
; GENERAL INFORMATION:  
; APPLICANT: INCYTE CORPORATION; TANG, Y. Tom;  
; APPLICANT: CORLEY, Neil C.; GUEGLER, Karl J.;  
; APPLICANT: BAUGHN, Mariah R.; LAL, Preeti G.;  
; APPLICANT: YUE, Henry; HILLMAN, Jennifer L.;  
; APPLICANT: AZIMZAI, Yalda  
; TITLE OF INVENTION: HUMAN TRANSFERASE PROTEINS  
; FILE REFERENCE: PF-0592-1 DIV  
; CURRENT APPLICATION NUMBER: US/10/427,631  
; CURRENT FILING DATE: 2003-04-29  
; PRIOR APPLICATION NUMBER: JS 09/786,240  
; PRIOR FILING DATE: 2002-03-12  
; PRIOR APPLICATION NUMBER: PCT/US99/20999  
; PRIOR FILING DATE: 1999-09-03

; PRIOR APPLICATION NUMBER: US 60/172,220  
; PRIOR FILING DATE: 1998-09-10  
; PRIOR APPLICATION NUMBER: US 60/155,248  
; PRIOR FILING DATE: 1998-11-04  
; PRIOR APPLICATION NUMBER: US 60/133,642  
; PRIOR FILING DATE: 1999-05-11  
; SOFTWARE: PERL Program  
; SEQ ID NO 33  
; LENGTH: 373  
; TYPE: PRT  
; ORGANISM: Caenorhabditis elegans  
; NAME/KEY: misc feature  
; OTHER INFORMATION: GenBank ID No. US20030175923A1 g2804432  
US-10-427-631-33

Query Match 56.1%; Score 1045; DB 12; Length 373;  
Best Local Similarity 55.3%; Pred. No. 1.1e-101;  
Matches 203; Conservative 64; Mismatches 92; Indels 8; Gaps 3;  
QY 1 MKALILVGGGTRLRPLTLSPKPLVDFAKPMILHQLAEKVEGVTEVVLAIRYRPEVM 60  
DB 9 MKALILVGGGTRLRPLTLTQPLVEFANKPMILHQLAEKVEGVTEVVLAIRYRPEVM 68  
QY 61 INFLKDFEDKLGITITCSQETPELGTAGPLALARDKLADGSGQPFVFLNSDVISYPPFAE 120  
DB 69 EQEMTHADRLGVKLIFSELEPELGTAGPLALARKHL-EGDA-PFFVFLNSDVICDFFPKQ 126  
QY 121 LIKFKHCHGGEATMTKVDPSKYGVVMEAEATGRVERFVEKPIFVGKINAGIYLLN 180  
DB 127 MVFEFKHKGEGTIAVTKVEPSKYGVVDFDQDKGIDDFVEKPIFVGKINAGIYIF 186  
QY 181 PSVLDRLELPTSLIEKEVFFQIAADQOLYAMVLPGFMDVQOPRDYITGLRLYLSIRKK 240  
DB 187 SKILDRIPLEKPTSLIEKEIFPEMFASGNLYAFVLPGFMDVQOPKDFLKGMSFLNHCHTT 246  
QY 241 SAAKLTG-----AHVGNVLHESAKIGEGCLIGDPDVAIGCGVVEDGVRLSRCTVMR 294  
DB 247 KSDKLETGSHPTATIRGNMVDPSATVGENCVIGDPDVGPRVKIEGGVRIHLHSTLS 306  
QY 295 GVRKKHACISNLIHGSTWGQWARIENMTILGEDVHVCDVYVNGGVLPKHKEIKSSI 354  
DB 307 DSSIGNVSWSGSVGRKCHIGSWRIENICVIGDDVVVMDVYVNGGVLPKHKEIKSSI 366  
QY 355 LKPEVM 361  
DB 367 PSKDIIM 373

RESULT 7  
US-10-128-714-3009  
; Sequence 3009, Application US/10128714  
; Publication No. US20030119013A1  
; GENERAL INFORMATION:  
; APPLICANT: Jiang, Bo  
; APPLICANT: Hu, Wenqi  
; APPLICANT: Tishkoff, Daniel  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Eroshkin, Alexey M  
; APPLICANT: Lemieux, Sebastien M  
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and  
; FILE OF INVENTION: Methods of Use  
; FILE REFERENCE: 10182-018-999  
; CURRENT APPLICATION NUMBER: US/10/128,714  
; CURRENT FILING DATE: 2002-04-23  
; PRIOR APPLICATION NUMBER: US 60/285,697  
; PRIOR FILING DATE: 2001-04-23  
; PRIOR APPLICATION NUMBER: US 60/287,066  
; PRIOR FILING DATE: 2001-04-27  
; PRIOR APPLICATION NUMBER: US 60/295,890  
; PRIOR FILING DATE: 2001-06-05  
; PRIOR APPLICATION NUMBER: US 60/303,899

; PRIOR FILING DATE: 2001-07-39  
; PRIOR APPLICATION NUMBER: US 60/316,362  
; PRIOR FILING DATE: 2001-06-31  
; NUMBER OF SEQ ID NOS: 8603  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 3009  
; LENGTH: 319  
; TYPE: PRT  
; ORGANISM: Aspergillus fumigatus  
US-10-128-714-3009

Query Match 47.1%; Score 578.5; DB 15; Length 319;  
Best Local Similarity 59.3%; Pred. No. 3.5e-84;  
Matches 170; Conservative 45; Mismatches 64; Indels 5; Gaps 3;  
QY 75 ITCQETPELGTAGPLALARDKLADGSGQPFVFLNSDVISYPPFAELIKFKHCHGGEATI 134  
DB 10 IEFVSESEPLGTAGPLKLA-EKILGKDDSPFVFLNSDIICDYPPKQLAEFFHKHGDCTI 68  
QY 135 MVTKVDEPSKYGVVME-EATGRVERFVEKPIFVGKINAGIYLLAPSVLDRLELRPTS 193  
DB 69 VVTKVDEPSKYGVVVKPNHPSRIDRFVEKPIFVGKINAGIYLLNPSVLKRIELRPTS 128  
QY 194 IEKEVFPQIADQOLYAMVLPGFMDVQOPRDYITGLRLYLSIRKKSAALKATGAHV- 252  
DB 129 IEQETFPALCSDGQLHSPDLGPFMDVQOPKDFLTGCLYLASLAKRNSKULAPNSPYV 188  
QY 253 --GNVLHESAKIGEGCLIGDPDVAIGCGVVEDGVRLSRCTVMRGVRIKKHACISNLIIG 310  
DB 199 YGVNVMVDPSAKIKNCRIEIPNVVIGPVVVDGVRQCVLLENSKYKDEAWIKSTIVG 248  
QY 311 WHSTVGQWARIENMTILGEDVHVCDVYVNGGVLPKHKEIKSSI 354  
DB 249 MNSSVGVKWARLENVTVLGDDVTIADVVYVNGGSILPHKSIKQNI 292

## RESULT 8

US-10-156-761-14505  
; Sequence 14505, Application US/10156761;  
; Publication No. US20030119018A2  
; GENERAL INFORMATION:  
; APPLICANT: OMURA, SATOSHI  
; APPLICANT: IKEDA, HARUO  
; APPLICANT: ISHIKAWA, JUN  
; APPLICANT: HORIKAWA, HIROSHI  
; APPLICANT: SHIBA, TADAYOSHI  
; APPLICANT: SAKAKI, YOSHIYUKI  
; APPLICANT: HATTORI, MASAHIRA  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-262  
; CURRENT APPLICATION NUMBER: US/10/-56,761;  
; CURRENT FILING DATE: 2002-05-29  
; PRIOR APPLICATION NUMBER: JP 2001-204089  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: JP 2001-272697  
; PRIOR FILING DATE: 2001-08-02  
; NUMBER OF SEQ ID NOS: 15109  
; SEQ ID NO 14505  
; LENGTH: 831  
; TYPE: PRT  
; ORGANISM: Streptomyces avermitilis  
US-10-156-761-14505

Query Match 27.6%; Score 515; DB 15; Length 831;  
Best Local Similarity 32.4%; Pred. No. 4.3e-45;  
Matches 121; Conservative 84; Mismatches 140; Indels 28; Gaps 11;  
QY 1 MKALILVGGGTRLRPLTLSPKPLVDFAKPMILHQLAEKVEGVTEVVLAIRYRPEVM 60  
DB 1 MKAVVMAGGEGTRLRPMTSSMPKPLPVNVRIMEHVRLRLKRLHGLNETVTVQFLASLV 60  
QY 61 INFLKDFEDKLGITITCSQETPELGTAGPLALARDKLADGSGQPFVFLNSDVISYPPFAE 120



Db 61 KNYFGDGE-LGMELTYANE3KPLGTAGSVKNAEALKDOTT---FLVISCDAITDFDLTE 116  
 Qy 121 LIKFHKGGEATIMVTKVDEPSKYGVVWMBEATGRVERFVEKPIFVGNKINAGIYLLN 180  
 Db 117 LINFHKEKAGALVTVCLTRVNPVLEFGITIVDE-EGKVERFLEKPTWQVP-SDTVNTGIY 174  
 Qy 178 LNPVLDRIELRPTSE--KEVFPQIAAD-QOLYAMVLPFGFMDVGPQPDYITG-----229  
 Db 175 VMEFEVNYVE-ADVSDWSGDVFPQMKDKPGYGVAGYWEVGTHTSVYKAAQAVL 233  
 Qy 230 ---LRLYLDISIRKSAAKLATGAHVGNVNLVHESAKIGEGCLIGPDV-----AIGFGCV 280  
 Db 234 EGKVDVEIDGFEISPGVWVAGAEVHPDAVLRGLYIGDYAKVEADVREHTVVGSNV 293  
 Qy 281 VEDGVRLSRCTVMRGVIRIKKHACISNIIGHSTVGOWARIENMTILGEDVHVDEVYGN 340  
 Db 294 VKSGAFLRHVVDVYVIGQHSNLRGCVIGKNTDIMAARIEDGAVIGDECLYGEESI 353  
 Qy 341 GGV-VLPHKEIKS 352  
 Db 354 GNVGVYFFKITEA 366

## RESULT 9

US-09-712-363-269  
 ; Sequence 269, Application US/097:2363  
 ; Patent No. US20020164588A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Eisenberg, David  
 ; APPLICANT: Rotstein, Sergio H.  
 ; APPLICANT: Marcotte, Edward M.  
 ; TITLE OF INVENTION: DETERMINING THE FUNCTIONS AND  
 ; FILE OF INVENTION: INTERACTIONS OF PROTEINS BY COMPARATIVE ANALYSIS  
 ; FILE REFERENCE: 07419-032001  
 ; CURRENT APPLICATION NUMBER: US/09/712.363  
 ; CURRENT FILING DATE: 2000-11-13  
 ; PRIOR APPLICATION NUMBER: PCT/US00/02246  
 ; PRIOR FILING DATE: 2000-01-28  
 ; PRIOR APPLICATION NUMBER: 60/179,531  
 ; PRIOR FILING DATE: 2000-02-01  
 ; PRIOR APPLICATION NUMBER: 60/117,844  
 ; PRIOR FILING DATE: 1999-01-29  
 ; PRIOR APPLICATION NUMBER: 60/118,206  
 ; PRIOR FILING DATE: 1999-02-01  
 ; PRIOR APPLICATION NUMBER: 60/124,593  
 ; PRIOR FILING DATE: 1999-03-26  
 ; PRIOR APPLICATION NUMBER: 60/134,093  
 ; PRIOR FILING DATE: 1999-05-14  
 ; PRIOR APPLICATION NUMBER: 60/134,092  
 ; PRIOR FILING DATE: 1999-05-14  
 ; PRIOR APPLICATION NUMBER: 60/165,124  
 ; PRIOR FILING DATE: 1999-11-12  
 ; PRIOR APPLICATION NUMBER: 60/165,086  
 ; PRIOR FILING DATE: 1999-11-12  
 ; NUMBER OF SEQ ID NOS: 292  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 269  
 ; LENGTH: 359  
 ; TYPE: PRT  
 ; ORGANISM: Mycobacterium tuberculosis  
 US-09-712-363-269

Query Match 26.2%; Score 488; DB 10; Length 359;  
 Best Local Similarity 34.6%; Pred. No. 8.3e-43;  
 Matches 125; Conservative 63; Mismatches 151; Indels 22; Gaps 10;  
 Qy 3 ALILVGGGTGRLRPLTLSPFPLVDFAKPMILHQIEALKEVGVTEVVLAINRPEVMIN 62  
 Db 8 AVLIVGGGTGRLRPLTLSPAPMLPTAGLPFLTLGRIAAAGIEHVILGTSYKPAV---64  
 Qy 63 FLKQPED--KLGITTCSETEPLCTAGPLALARDKLDGSGQPFVNLSDVISEYPAE 120  
 Db 65 FEAEFGDGSALGLQIEYVTEEHPGLTGGGIANVAGKLRNDTA---MVFNGBVLSGADLAQ 121

Qy 121 LIKFHKGGEATIMVTKVDEPSKYGVVWMBEATGRVERFVEKPIFVGNKINAGIYLLN 180  
 Db 122 LDFHRSNRADVTLLQVVRVGDPRAFGCVPTDE-EDRVVAFLEKTEDEPTDQINAGCVFE 180  
 Qy 181 PSVLDRI-ELRPTSEIEKEVFPQIAADQ--OLYAMVLPFGFMDVGPQPDYITGLRLYLD 237  
 Db 181 RNVIDRIFQGEVSEVEREFPALLADGDCIKYGVVDASYWRDMGTPEDFVRG---SADLV 237  
 Qy 238 RKSAAKLATGAHVGNVNLVHESAKIGEGCLIGPDVAGFCCVVEDGVRLSRTCVRMGVR 297  
 Db 238 RGIAPSPALRGHR--GEQLVHDGAASPGALLIGGTVWGRGAEIGPGLTRLCGAVIFDGV 295  
 Qy 298 IKKHACISNIIGHSTVGOWARIENMTI--IGEDVHVDEVYGN---GVVLPHKEIKS 352  
 Db 296 VEAGCVIERSIIGFGARIGPRLIRCGVIGDGADIGARCELLSGARVMPGVFLPDGGIRY 355  
 Qy 353 \$ 353  
 Db 356 \$ 356

## RESULT 10

US-10-156-761-12571  
 ; Sequence 12571, Application US/10:56761  
 ; Publication No. US20030119018A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: OMURA, SATOSHI  
 ; APPLICANT: IKEDA, HARUO  
 ; APPLICANT: ISHIKAWA, JUN  
 ; APPLICANT: HORIKAWA, HIROSHI  
 ; APPLICANT: SHIBA, TADAYOSHI  
 ; APPLICANT: SAKAKI, YOSHIYUKI  
 ; APPLICANT: HATTORI, MASAEIRA  
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
 ; FILE REFERENCE: 249-262  
 ; CURRENT APPLICATION NUMBER: US/10/156.761  
 ; CURRENT FILING DATE: 2002-05-29  
 ; PRIOR APPLICATION NUMBER: JP 2001-204089  
 ; PRIOR FILING DATE: 2001-05-30  
 ; PRIOR APPLICATION NUMBER: JP 2001-272697  
 ; PRIOR FILING DATE: 2001-08-02  
 ; NUMBER OF SEQ ID NOS: 15109  
 ; SEQ ID NO 12571  
 ; LENGTH: 360  
 ; TYPE: PRT  
 ; ORGANISM: Streptomyces avermitilis  
 US-10-156-761-12571

Query Match 25.2%; Score 469.5; DB 15; Length 360;  
 Best Local Similarity 32.9%; Pred. No. 7.5e-41;  
 Matches 114; Conservative 68; Mismatches 147; Indels 17; Gaps 8;  
 Qy 2 KALILVGGGTGRLRPLTLSPFPLVDFAKPMILHQIEALKEVGVTEVVLAINRPEVMI 61  
 Db 3 EAILVGGGTGRLRPLTVHTKPMVPAAGVPLFTHQLARARAAAGVDHVLATSYLA 62  
 Qy 62 NFLKDFEDKLGITTCSETEPLCTAGPLALARDKLDGSGQPFVNLSDVISEYPAEL 121  
 Db 63 PYFGD-GSALGLHLEYVTEEPLGTGGAINRVARLSHSGDPEVLIFNGDILTGDLRAL 121  
 Qy 122 IKFHKHKGGEATIMVTKVDEPSKYGVVWMBEATGRVERFVEKPKI---FVGNKINAGIYL 178  
 Db 122 VHTHETTAADVSLHLLTKVTDPRAYGLVPTDE-TGRVTAFLKPKPTPEEIVTDQINAGAYV 180  
 Qy 179 LNPVLDRIEL-RPTSEIEKEVFP-QIAADQOLYAMVLPFGFMDVGPQPDYITGLRLYLD 236  
 Db 181 FRRSVIDTIPAPRPVSVERETFPGLLSAGAHLOGKVDSTYWLDLGTFAAFVRG---SADL 237  
 Qy 237 IRKSAAKLATGAHVGNVNLVHESAKIGEGCLIGPDVAGFCCVVEDGVRLSRTCVRMGV 296  
 Db 238 VLGRAPSPAVPGR--CGDRJVLPTASVANSDAKLTGGTVGEGAFVGEGARIFGSTLLSGA 295

```
QY 297 RIKKHACISNIIQWHSITVGQ-----WARIENMTILGEDVHVCDV 337
Db 296 VWFPGAVITSLGARSVGRSILTGAVIGDGAVIDAGNELRDGV 341

RESULT 11
US-09-738-626-4325
; Sequence 4325, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, WASATO
; APPLICANT: OKAZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738.626
; CURRENT FILING DATE: 2000-12-18
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: JP 00/280988
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: Patent in ver. 3.0
; SEQ ID NO 4325
; LENGTH: 348
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-4325

Query Match 23.1%; Score 431.5; DB 10; Length 348;
Best Local Similarity 31.7%; Pred. No. 7.5e-37;
Matches 113; Conservative 66; Mismatches 158; Indels 19; Gaps 9;

QY 6 LVGFGTRLRPLTSLSPKPLVDANKPMILHQIEALKVEGVTEVLAINYRPEVMINFLK 65
Db 1 MVGGKTRLRPLTNTPKMLPTAGHFFLTHLARIKAAGITHWLGTSPKAEVFEYFG 60

QY 66 DFEDKLGITITCSQETEPGLTAGLALARDKLADGSGQPFVLSNDVISEYPPAEILKPH 125
Db 61 D-GSEMGLEIYYVVEDQPLGTGGIRNVYDKLRHDTA--IVFNGDVLGADLNSILDTH 116

QY 126 KCHGGEATIMTKVDEPSKYGVVMEETATGVRFEVKPKIFVGNKINAGIYLLNPSVLD 185
Db 117 REXDADLTMLHLVAVNAPRASFSCVPTDE-DGRVSEFLEKTEDPTDQINAGCYVFKKELIE 175

QY 186 RIEL-REPTSKEKEVFPQIAAD-QOLYAMVLPGFMDVGPDRDVI TGLRLYLDLSIRKKSAA 243
Db 176 QIPAGRAVVERETFPOLLEGGKRVGHVDASTWRDMGTSPDFRG---SADLVRGIAYS 232

QY 244 KLATGAHVGNVNLVHESAKIEGCLIGDPVAIGPGCVVEDGVRLSCTVMRGVRIKKHAC 303
Db 233 PLLEGK--TGESLVDSAGV3DGVLLGGTVVGRGTGAGCRVDNTVIFDGVITIEPGAV 230

QY 304 ISNSIIQWHSITVGWARIENMTILGEDVHVCDVYNSG-----GVVLPKHEIKSS 353
Db 291 IENSIISGARIGANAHISG-CIIGEAQVGARCELNAGMRVFGVVPDPSGIRFS 345

RESULT 12
US-09-769-734-33
; Sequence 33, Application US/09769734
; Publication No. US20030143666A1
; GENERAL INFORMATION:
; APPLICANT: Ecopia BioSciences Inc.
```

```
; TITLE OF INVENTION: Genetic Locus for Everninomicin Biosynthesis
; FILE REFERENCE: PA 005-US
; CURRENT APPLICATION NUMBER: US/09/769.734
; CURRENT FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 33
; LENGTH: 355
; TYPE: PRT
; ORGANISM: M. carbonacea
US-09-769-734-33
```

```
Query Match 16.8%; Score 314; DB 12; Length 355;
Best Local Similarity 27.2%; Pred. No. 2e-24;
Matches 100; Conservative 68; Mismatches 148; Indels 52; Gaps 12;

QY 1 MKALILVGGFGTRLRPLTSLSPKPLVDANKPMILHQIEALKVEGVTEVLAINYRPEVM 60
Db 1 VKALVLAGGIGSRMRPITHTSKQJIPVANKPVLFPYGJEARIDAGIREVGIIVGSTAPEI 60

QY 61 INFLKDFEDKLGITITCSQETEPGLTAGLALARDKLADGSGQPF----FVLN--SDVIS 114
Db 61 ERAVGD-GSQGLKVYLYLPQDAPRGLGHAVLIARDFLCDDBFVWYLGDNFVLGGINDAVE 119

QY 115 ----EYPPAEILKHKHKGGEATIMTKVDEPSKYGVVMEETATGVRFEVKPKIFVGN 170
Db 120 RFRERPHAOQ-----MLTKVKDPHAFGIATM-GPDGRVVDVVEKPRYPKSD 165

QY 171 KINAGIYLLNPSVLDRI-ELRPT-----SIEKEVFPQIAADQOLYAMVLPGFMDVGPDR 224
Db 166 LALGVYVFPVVEHAIKELKPSWRNELEITDAQWLHDHRRRESTITGFWKDTGSLA 225

QY 225 DYTGLRLYLDLSIRKKSAAKLATGAHVGNVNLVHESAKIEGCLIGDPVAIGPG----- 278
Db 226 DMJEMNRFILSLDSEVSGVSADTEITGRVVIGPGAVITGSRITGP-VVVGAGSIIRNS 284

QY 279 -----CVVEDGVLSRCTVMRGVRIKKHACISNIIQWHS--TVGOWARIENMT 325
Db 285 QIGPFTSIDCCTVIDS-EIEQSIVLRGAFIDG-GRIEWSMIGREARLTGPRAPKTYRF 343

QY 326 ILGEDVHV 333
Db 344 VLGDHSEV 351
```

```
RESULT 13
US-09-922-683-10
; Sequence 10, Application US/09922683
; Publication No. US20020192793A1
; GENERAL INFORMATION:
; APPLICANT: DECKER, Heinrich
; TITLE OF INVENTION: ISOLATION OF THE BIOSYNTHESIS GENES FOR
PSEUDO-CLIGOSACCHARIDES FROM STREPTOMYCES GLAUCESCENS
GLA.O AND THEIR USE
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
ADDRESS: POLEY & LARDNER
STREET: 3000 K Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20007-5109
; COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/922,683
FILING DATE: 07-Aug-2001
CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/194,905
```

FILING DATE: 1999-12-01  
APPLICATION NUMBER: DE 19622783.6  
FILING DATE: 07-JUN-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Granados, Patricia D.  
REGISTRATION NUMBER: 33,683  
REFERENCE/DOCKET NUMBER: 026083/0193  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 672-5300  
TELEFAX: (202) 672-5399  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 355 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 10:  
US-09-922-683-10

Query Match 15.9%; Score 296; DB 10; Length 355;  
Best Local Similarity 27.5%; Pred. No. 1.6e-22;  
Matches 95; Conservative 58; Mismatches 160; Indels 32; Gaps 9;

QY 1 MKALILVGGFGTGLRPLTLSPKPLVDFAFKPMILHQIEALKEVGVTEVVVLAIN-YRPEV 59  
DB 1 VKALVLAGGTGSLRPLTPHTAAKQLPIANKPVLFYALESLAAGVREAGVVVGAYCRE- 59  
QY 60 MINFLKDFEDKLGITITCSQETPLGTAGPLALARDKLAGSGQPPFVLNSOVISEYPPA 119  
DB 60 -IRELTGDTGFAFLIRITLHQPRPLGLAHAVRIARGFLGD---DDFLYLGDNYLPQGV 115  
QY 120 ELIKFKHCHGGEATIMVTKVDEPSKYGVVMEATGRVERFVEKPKIFVGNKINAGIYLL 179  
DB 116 DFAQSAADPAARALLTPVADPSAFGAEV-DADGNVLELEKPDVPRSSJALIGVYAF 174  
QY 180 NPSVLDRIE-LRPTS-----EKEVFQIAADQOLYAMVLPGFMDVGQPRDYITGLRLY 233  
DB 175 SPAVHEAVRAITPSARGELE-THAVQWMDRLGLRVRAETTRPRDTGSAEDMLEVNRHV 234  
QY 234 LDSIRKSAK-----LATGAHVGNVLVSHESAKIGEGCLIGDPVAIG 276  
DB 235 LOGLEGRIEGKVDASHTLVGRVRAEGAIVRGSHVGVVIGAGAVVSNSS-VGPVTSIG 293  
QY 277 PGCVVGVGLRSLCTVMRGVRIKKHACISNIIGHSHVTCQWARI 321  
DB 294 EDCRVESA-IEYSLVRGAQVGAASRIEASLIGRGAVVGAPRL 337

RESULT 14  
US-10-156-761-8552  
Sequence 8552, Application US/10156761  
Publication No. US20030119018A1  
GENERAL INFORMATION:  
APPLICANT: OMURA, SATOSHI  
APPLICANT: IKEDA, HARUO  
APPLICANT: IGHIKAWA, JUN  
APPLICANT: HORIKAWA, HIROSHI  
APPLICANT: SHIBA, TADAYOSHI  
APPLICANT: SAKAKI, YOSHIYUKI  
APPLICANT: HATTORI, MASAHISA  
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
FILE REFERENCE: 249-262  
CURRENT APPLICATION NUMBER: US/0156,761  
CURRENT FILING DATE: 2002-05-29  
PRIOR APPLICATION NUMBER: JP 2001-204089  
PRIOR FILING DATE: 2001-05-30  
PRIOR APPLICATION NUMBER: JP 2001-272697  
PRIOR FILING DATE: 2001-08-02  
NUMBER OF SEQ ID NOS: 15109  
SEQ ID NO 8552  
LENGTH: 237  
TYPE: PRT

ORGANISM: Streptomyces avermitilis  
US-10-156-761-8552

Query Match 15.1%; Score 281.5; DB 15; Length 237;  
Best Local Similarity 30.7%; Pred. No. 3e-21;  
Matches 70; Conservative 46; Mismatches 101; Indels 11; Gaps 6;

QY 1 MKALILVGGFGTGLRPLTLSPKPLVDFAFKPMILHQIEALKEVGVTEVVVLAINYRPEV 59  
DB 1 KHAVILAGGKGRVLRPYTTALPKPLVPIGDQHALEIVLRQLAAGAGFTGCTIA:GHLGEI 60  
QY 60 MINFLKDFEDKLGITITCSQETPLGTAGPLALARDKLAGSGQPPFVLNSDVISEYPPA 119  
DB 61 IRAYVGD-GSQMGAVDYATESPLGTMGPLLTNRERLPES----FLVMNGDILTLDDYA 115  
QY 120 ELIKFKHCHGGEATIMVTKVDEPSKYGVVMEATGRVERFVEKPKIFVGNKINAGIYLL 179  
DB 116 DVLQRDRSGAPLTIATYARKVHIDFGVLTQ--AGRVVGFTEKPS--MDYRVSMGVYGL 171  
QY 180 NPSVLDRIEL-RPTSIEKEVFQIAADQOLYAMVLPGFMDVGOPROY 226  
DB 172 SRKTLGGYTPGLGFDLVLLELLKTNLPHAYDFDGYWLDIGRPDDY 219

RESULT 15  
US-09-925-637-18  
Sequence 18, Application US/09925637  
Patent No. US20020103338A1  
GENERAL INFORMATION:  
APPLICANT: Choi  
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Polypeptides  
FILE REFERENCE: PB560  
CURRENT APPLICATION NUMBER: US/09/925,637  
CURRENT FILING DATE: 2001-08-10  
PRIOR APPLICATION NUMBER: PCT/US00/23773  
PRIOR FILING DATE: 2000-08-31  
PRIOR APPLICATION NUMBER: US 60/151,933  
PRIOR FILING DATE: 1999-09-01  
PRIOR APPLICATION NUMBER: US 08/781,966  
PRIOR FILING DATE: 1997-01-03  
PRIOR APPLICATION NUMBER: US 08/956,171  
PRIOR FILING DATE: 1997-10-20  
PRIOR APPLICATION NUMBER: US 60/009,861  
PRIOR FILING DATE: 1996-01-06  
NUMBER OF SEQ ID NOS: 74  
SOFTWARE: Patentin ver. 2.3  
SEQ ID NO 18  
LENGTH: 453  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-925-637-18

Query Match 14.9%; Score 277.5; DB 10; Length 453;  
Best Local Similarity 27.2%; Pred. No. 2.2e-20;  
Matches 106; Conservative 64; Mismatches 146; Indels 73; Gaps 15;

QY 3 ALILVGGFGTGLRPLTLSPKPLVDFAFKPMILHQIEALKEVGVTEVVVLAINYRPEVWIN 62  
DB 8 AILLAAGKGRTRKSKY---KVLHEVAGKPMVHVLESVKSGVDQVVVTIYGHGAE---- 60  
QY 63 FLKDFEDKLGITITCSQETPLGTAGPLALARDKLAGSGQPPFVL-NSDVISEYPPFAEL 121  
DB 61 ---SVKGHGERSLSYFQEEQLGTAAHVQMAKSHLEDEKGTIVVCGDTPLITKETLVL 117  
QY 122 IKFKHCHGGEATIMVTKVDEPSKYGVVMEATGRVERFVEKPKIFVG-----NKNAGIY 177  
DB 118 IAHEDANAQAATVLSASTQQPYGYGRIV-RNASGRLEIRIVEEKDATCAEKDINEISSGIF 176  
QY 178 LNPVSLDRIELRPTSIEKEVFQIAADQOLYAMVLP-----GFMDVGOPROY-- 226  
DB 177 AFNNKTL-----FEKLTQVQNDNAQGEYLPDLVLSLINDGGIVEYVTRNDVEE 225  
QY 227 ITGL--RLYLDISIRKKSAAKLATGAHVGVGNVLVHESAKIGEGCLIGDPVAIGSPCVVEDG 284

Search completed: October 9, 2003, 14:08:10  
Job time : 32 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 9, 2003, 13:59:16 ; Search time 20 Seconds  
(without alignments)  
763.711 Million cell updates/sec

Title: US-09-374-967-2

Perfect score: 1364  
Sequence: 1 MKALILVGGFGTRLRP-TLS.....GVVLPHEKISILKPEIWM 361

Scoring table: BLOSUM62  
Gapop 10.0 , Gapex: 0.5

Searched: 3287.7 seqs, 42310358 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1681	90.2	361	4	US-09-786-240-32
2	1183.5	63.5	360	4	US-09-786-240-2
3	1178	63.2	361	3	US-09-032-372-12
4	1045	56.1	373	4	US-09-786-240-33
5	401	21.5	421	3	US-09-032-372-3
6	307	16.5	355	1	US-08-196-218-34
7	307	16.5	355	1	US-08-681-953-34
8	296	15.9	355	4	US-09-194-905-10
9	261	14.0	233	4	US-09-328-352-7812
10	241.5	13.0	461	4	US-09-134-001C-3316
11	239	12.8	146	4	US-09-370-838-190
12	228.5	12.3	224	4	US-09-199-637A-175
13	227.5	12.2	258	4	US-09-252-991A-17377
14	214	11.5	393	4	US-09-107-532A-6010
15	193	10.4	431	1	US-08-469-202-14
16	193	10.4	431	2	US-08-484-434C-14
17	193	10.4	431	4	US-09-384-361-14
18	191	10.2	431	1	US-08-090-523-4
19	191	10.2	431	1	US-08-398-627-4
20	191	10.2	431	1	US-08-406-858-4
21	191	10.2	431	4	US-08-120-703A-4
22	191	10.2	431	4	US-08-399-023-4
23	191	10.2	431	5	PCT-US91-04036-4
24	191	10.2	431	5	PCT-US94-05275-4
25	191	10.2	488	4	US-09-444-728-2
26	190	10.2	292	3	US-09-320-878-14
27	190	10.2	292	3	US-09-105-537-12

28	190	10.2	292	4	US-09-141-908-16	Sequence 16, Appl
29	190	10.2	292	4	US-09-657-440-14	Sequence 14, Appl
30	190	10.2	3782	3	US-09-105-537-4	Sequence 4, Appl
31	188	10.1	431	1	US-08-090-523-2	Sequence 2, Appl
32	188	10.1	431	1	US-08-398-627-2	Sequence 2, Appl
33	188	10.1	431	1	US-08-406-858-2	Sequence 2, Appl
34	188	10.1	431	4	US-08-120-703A-2	Sequence 2, Appl
35	188	10.1	431	4	US-08-399-023-2	Sequence 2, Appl
36	188	10.1	431	5	PCT-US91-04036-2	Sequence 2, Appl
37	188	10.1	431	5	PCT-US94-05275-2	Sequence 3, Appl
38	186.5	10.0	518	1	US-08-485-241-3	Sequence 3, Appl
39	186.5	10.0	518	2	US-08-874-162-3	Sequence 5, Appl
40	186	10.0	517	1	US-08-485-241-5	Sequence 5, Appl
41	186	10.0	517	2	US-08-874-162-5	Sequence 5, Appl
42	186	10.0	518	1	US-08-299-675-2	Sequence 2, Appl
43	185	9.9	461	4	US-09-328-352-5107	Sequence 5107, Ap
44	183.5	9.8	324	4	US-09-107-532A-4762	Sequence 4762, Ap
45	182.5	9.8	296	4	US-09-107-532A-6365	Sequence 6365, Ap

ALIGNMENTS

RESULT 1  
US-09-786-240-32  
; Sequence 32, Application: US/39786240  
; Patent No. 6558935

; GENERAL INFORMATION:  
; APPLICANT: INCYTE PHARMACEUTICALS, INC.  
; APPLICANT: TANG, Y. Tcm  
; APPLICANT: CORLEY, Neil C.  
; APPLICANT: GUESLER, Karl J.  
; APPLICANT: BAUGHN, Mariah R.  
; APPLICANT: LAU, Preeti  
; APPLICANT: YUE, Henry  
; APPLICANT: HILLMAN, Jennifer L.  
; APPLICANT: AZIMZAI, Valda

; TITLE OF INVENTION: HUMAN TRANSFERASE PROTEINS

; FILE REFERENCE: PF-0592 PCT  
; CURRENT APPLICATION NUMBER: US/09/786,240

; CURRENT FILING DATE: 2002-03-12

; PRIOR APPLICATION NUMBER: 09/150,657; Unassigned; 09/186,779; Unassigned; 60/123,642

; PRIOR FILING DATE: 1998-09-10; 1998-09-10; 1998-11-04; 1998-11-04; 1999-05-11

; NUMBER OF SEQ ID NOS: 33

; SOFTWARE: PERL Program

; SEQ ID NO 32

; LENGTH: 361

; TYPE: PRT

; ORGANISM: Arabidopsis thaliana

; PUBLICATION INFORMATION:

; DATABASE ACCESSION NUMBER: g2642159

US-09-786-240-32

Query Match 90.2%; Score 1681; DB 4; Length 361;

Best Local Similarity 88.6%; Pred. No. 6.1e-179;

Matches 320; Conservative 22; Mismatches 19; Indels 0; Gaps 0;

QY 1 MKALILVGGFGTRLRPLTLSPFKPLVDVDFANKPMTLHQIEALKVEGVTVTVLAINYREVM 60

DB 1 MKALILVGGFGTRLRPLTLSPFKPLVDVDFANKPMTLHQIEALKVGVDEVVLAINYQPEVM 60

QY 61 INFLKDFEDKLGITITCSQTEPLGTAGPLALARDKLADGSGQPFVFLNSDVISEYPPFAE 120

DB 61 LNFLKDFETKLEIKTCQSQTEPLGTAGPLALARDKLADGSGEPFVFLNSDVISEYPLKE 120

QY 121 LKFKHCHGGEGATIMVTKVDEPSKYGVVVMEEATGRVERFVEKPKIFVGNKINAGIYLLN 180

DB 121 MLEFHKSHGGEASIMVTKVDEPSKYGVVVMEEATGRVEKPKIYVGNKINAGIYLLN 180

QY 181 PSVLDRIELRPTSTEKEVFPQIAAQQLYANVLGFWMDVQCPRDYITGLRLYLDLSIRKK 240

DB 181 PSVLDRIELRPTSTEKEVFPKIAAQQLYANVLGFWMDVQCPRDYITGLRLYLDLSIRKK 240



Db 180 NPEVIDLIEMKPTSEIKETPILVEEKQLYFDFLEGFMVDVGQPKDFLSGTVLNLSLAK 239  
Qy 240 KSAAKLATGARHVGNVLVHESAKIGECGLIGPDVAIGPGCVGVEDGVRLSRCTVMRGVRIK 299  
Db 240 RQPKLATGANIVGNALIDPTAKISSTAKIGPDVIGPNVTIGDGVRIITRSVVLNCSNLIK 299  
Qy 300 KHACISNIIIGMHSITVGOWARIEIENMTILGEDVHVGVDEVSNGGVVLPKHKEIKSILKPEI 359  
Db 300 NLSLVKSTIVGWNSTVGWCLEGTVLGDDEVEYKDBIYNGGKVLPHKSIQSNVPEKAI 359  
Qy 360 VM 361  
Db 360 IM 361

RESULT 4  
US-09-786-240-33  
; Sequence 33, Application US/09786240  
; Patent No. 6558935  
; GENERAL INFORMATION:  
; APPLICANT: INCYTE PHARMACEUTICALS, INC.  
; APPLICANT: TANG, Y. Tom  
; APPLICANT: CORLEY, Neil C.  
; APPLICANT: GUEGLER, Karl J.  
; APPLICANT: BAUGHN, Mariah R.  
; APPLICANT: LAL, Preeti  
; APPLICANT: YUE, Henry  
; APPLICANT: HILLMAN, Jennifer L.  
; APPLICANT: AZIMZAI, Yalda  
; TITLE OF INVENTION: HUMAN TRANSFERASE PROTEINS  
; FILE REFERENCE: PF-0592 PCT  
; CURRENT APPLICATION NUMBER: US/09/786,240  
; PRIOR FILING DATE: 2002-03-12  
; PRIOR APPLICATION NUMBER: 09/150,657; unassigned; 09/186,779; unassigned; 6C/133,642  
; PRIOR FILING DATE: 1998-09-10; 1998-09-10; 1998-11-04; 1998-11-04; 1999-05-11  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: PERL Program  
; SEQ ID NO 33  
; LENGTH: 373  
; TYPE: PRT  
; ORGANISM: Caenorhabditis elegans  
; PUBLICATION INFORMATION:  
; DATABASE ACCESSION NUMBER: g2804432  
US-09-786-240-33

Query Match 56.1%; Score 1045; DB 4; Length 373;  
Best Local Similarity 55.3%; Pred. No. 6.3e-108;  
Matches 203; Conservative 64; Mismatches 92; Indels 8; Gaps 3;

Qy 1 MKALILVGGFGRTRLPRLTSLSPKPLVDFPANKPMILHQIEALKEVGVTEVVLAINYRPEVM 60  
Db 9 MKALILVGGYGRTRLPRLTLTPKPLVEFANKPMILHQIEALAEVGVDTVVLVSYRAEQL 68  
Qy 61 INFLKDFEDKLGITITCSQETEPGLTAGPLALARDKLADGSGQPPFVLSNDSVISEYPAE 120  
Db 69 EQEMTVHADRLGVKLIFSLSEPLGLTAGPLALARKHL-EGDA-PFFVLSNDSVICTDPPFKQ 126  
Qy 121 LKIFKHCHGGGATIMVTWKDPSKYGVVMEETGRVERFVEKPKIFVGNKINAGIYLLN 180  
Db 127 MVEFHNGKGGTATVTKVEEPSKYGVVVFQDDGKIDDDFVEKQEVYVGNKINAGLIYFS 186  
Qy 181 PSVLDRIELRPTSTIEKEVFPFOAADQQLYAMVLPGFWMVDGQPRDYITGLRLYLDSTIRK 240  
Db 187 SKILDRIPLKPTSTIEKEIFPEMAFSGNLVYAVLPGFWMVDGQPKDFLKGMSLPLNHCHTT 246  
Qy 241 SAAKLATG-----AHVGNVLVHESAKIGEGCLIGPDVAIGPCVVDGVRLSRCTVMR 294  
Db 247 KSKLETGSIHTPATIRGNTWVDP SATVGENCVIGPDVWIGPRVKIEGGVRIILHSTLS 306  
Qy 295 GVRIKKHACISNIIIGMHSITVGOWARIEIENMTILGEDVHVGVDEVSNGGVVLPKHKEIKSI 354  
Db 307 DSSIGNYSWGSIVGRKCHIGSWVRRIENICVIGDGVVVKDEYINGASVLPKSIAYNV 366

Qy 355 LKPEIVM 361  
Db 367 PSKOIIM 373

RESULT 5  
US-09-032-372-3  
; Sequence 3, Application US/09032372  
; Patent No. 6008337  
; GENERAL INFORMATION:  
; APPLICANT: Bandman, Olga  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Guegler, Karl J.  
; APPLICANT: Yue, Henry  
; APPLICANT: Lal, Preeti  
; TITLE OF INVENTION: CELL CYCLE RELATED PROTEINS  
; NUMBER OF SEQUENCES: 13  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/032,372  
; FILING DATE: Herewith  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0478 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650-855-0555  
; TELEFAX: 650-845-4166  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 421 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: COLNROT23  
; CLONE: 1693222  
US-09-032-372-3

Query Match 21.5%; Score 401; DB 3; Length 421;  
Best Local Similarity 29.1%; Pred. No. 5.8e-36;  
Matches 123; Conservative 61; Mismatches 159; Indels 80; Gaps 13;

Qy 1 MKALILVGG--FGTRLRPLTSLSPKPLVDFPANKPMILHQIEALKEV-GVTEVVLAINYRP 57  
Db 2 LKAVILIGQPKGTRFRPLSPFVPEKPLFPVAGVPMIQHHIEACAQVPMQSIILLIGFYQP 61  
Qy 58 -EVMINFLKDFEDKLGITITCSQETEPGLTAGPLALARDKLADGSGQPPFVLSNDSVISEY 116  
Db 62 DBPLTFQFEAAQQEENFLPVRYLQEFAPLGTGGGLYHFRDQLAGSPFAFFVVLNADVCSD 121  
Qy 117 PPAELIKFKHCHGGGATIMVTWKD--BPSKYGVVMEETGRVERFVEKPKIFVGNKINA 174  
Db 122 PLSANLEAHRQRPFJLGTITANTQTSLNYGCVIENPQTHEVLHYVEKSTFISDINC 181  
Qy 175 GYLLNPSVL-----DRIELRP-----TSIEKEVFTQIAADQQLYAVVL 213

```
Db 182 GIYLFSPALKPLRDVFNQDQLEDSPGLPGAGTIRLEQDVFSALAGQCIYVHLT 241
QY 214 PGFMDVQOPRDYITGLRLYLDSIRKSAAKLATGAHVGVNVL-----VHESAKIGEGCL 268
Db 242 DGIWSQISAGSALYSRLYRVDTHPERLA--KHTPGMPDPRECVPHPDROG----- 295
QY 269 IGPDAIAGP-----GCVBDEGYRLSRC--TVNRGVRIKKHACISNSIIGHSTVG 316
Db 296 -GPLGCAGPQRLHREGDGRG-RVCGSGRASSMPELCKRSTRV---FCIASWAGAPMDAG 351
QY 317 QWARIENWTL-----LGEDVHVUCDEYVSGGVLPKHKEI 350
Db 352 RAWRVPPVTLPTTPEPAWTVRASSRTGSCCLLSPSMGCRVRIPAEVLILNSIVLPKHKEI 411
QY 351 KSS 353
Db 412 SRS 414

RESULT 6
US-08-196-218-34
; Sequence 34, Application US/08195218
; Patent No. 5614619
; GENERAL INFORMATION:
; APPLICANT: Piepersberg, Wolfgang
; APPLICANT: Stockmann, Michael
; APPLICANT: Taleghani, Kampiz Mansouri
; APPLICANT: Distler, Jurgen
; APPLICANT: Grabley, Susanne
; APPLICANT: Sichel, Petra
; APPLICANT: Brau, Barbara
; TITLE OF INVENTION: Secondary-Metabolite Biosynthesis Genes
; TITLE OF INVENTION: From Actinomycetes, Method of Isolating Them, and Their
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: United States
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 25-AUG-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ogden, Stasia L.
; REGISTRATION NUMBER: 36,228
; REFERENCE/DOCKET NUMBER: 02481.1372-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4400
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 355 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-196-218-34

Query Match 16.5%; Score 307; DB 1; Length 355;
Best Local Similarity 28.4%; Pred. No. 1.4e-25;
Matches 100; Conservative 59; Mismatches 137; Indels 56; Gaps 12;

QY 1 MKALIVLAGSGTRLRPSYSMPKQLIPFIANTPVLHVHNAVRELGVTEVSVIVGNRGPEI 60
Db 60 MINFLKDFEDKLG-----TITCSQETELGTAGPLALARDKLADGSGQPFFVFLNSDVI 113
Db 61 -----EAVLGGAGFADVRITTFQAPRGLAHTVSIAGFJGD---DDFVWYLGDNM 109
QY 114 SEYPFAELIKFKHCHGSEATINVTIKVDEPSKYGVVVMEEATGRVERFEVKPKIFVGNKIN 173
Db 110 LPDGVTEIAEFTQRPAACVVVHKVPDPRSGVAEL-GPDGEVLRLVEKFWQPRSDMAL 168
QY 174 AGIYLNDSVLDRI-EARPTS-----IEKEVFPQIADQQLYANVLPGFMDVQOPDIYI 227
Db 169 IGWVFFTAIHQAAVAAISPSRSGELETIDAVQWLVTSQADVRLSYDGYWKTGRVEDVL 228
QY 228 TGLRLYLDSIRKSAAKLATGAHVGVNVLVHESAK-----ICEGCL-----IGP 271
Db 229 ECNSHLLDGLTPRVGGQVADSVLGVVVIEAGARIVRSRVEGPAIIGAGTVLQDSQVGP 288
QY 272 DVAIGPGCVVED-----GVRLSRCTV--MRCVRIKKHACISNSIIGHSTVG 316
Db 289 HTSIGRDCVTDSRLSGSIALDEASVTQVRGLR-----NSLIGRAASVVG 332

RESULT 7
US-08-681-953-34
; Sequence 34, Application US/08681953
; Patent No. 5710032
; GENERAL INFORMATION:
; APPLICANT: Piepersberg, Wolfgang
; APPLICANT: Stockmann, Michael
; APPLICANT: Taleghani, Kampiz Mansouri
; APPLICANT: Distler, Jurgen
; APPLICANT: Grabley, Susanne
; APPLICANT: Sichel, Petra
; APPLICANT: Brau, Barbara
; TITLE OF INVENTION: Secondary-Metabolite Biosynthesis Genes
; TITLE OF INVENTION: From Actinomycetes, Method of Isolating Them, and Their
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: United States
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 30-JUL-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/196,218
; FILING DATE: 25-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Ogden, Stasia L.
; REGISTRATION NUMBER: 36,228
; REFERENCE/DOCKET NUMBER: 02481.1372-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4400
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 355 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
```





```
RESULT 10
US-09-134-001C-3316
; Sequence 3316, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3316
; LENGTH: 461
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3316

Query Match      13.0%; Score 241.5; DB 4; Length 461;
Best Local Similarity 25.7%; Pred. No. 4.2e-18;
Matches 98; Conservative 66; Mismatches 159; Indels 59; Gaps 15;

Qy 3 ALILVGGFGRRLRLTLSPKPLVDFAKPMILHQIEALKEVGVTEVVLAINRPEVMIN 62
Db 15 AIIAAGKGRMKSKY---KVLHEVAGKPWEHVLNNVKQAGVDQIVTIIGHGAE--- 67

Qy 63 FLKDFEDKLGITITCQETEPGLTAGPLALARDKLAGSQPFVL-NSDVISPEPAEL 121
Db 68 ---SVKDTLGNQSLYSFODKSLGTAHAYKMAHEHLADKEGT-LVVCGDTPLTITQTLSL 124

Qy 122 IKPHKCHGGEATIMVTKVDESKYGVVMEERATGRVERFVEKPIFVGK---INAGIV 177
Db 125 IEHSTQSHVTVLSASTINPYGGRIRNTH-NGILRIVEEKDANDSERAIKEISSGIP 183

Qy 178 LLNRSVL-DRIELRPTSTEKEVFOIAADQQLYAMV-PGFMDVGGPRDYIT-----G 229
Db 184 AFNRVLPEKLE-----QVNDNAQGEYVLPOVLSLILKDGKAEVYCTEDFDEIIG 235

Qy 230 L--RLYLDLSIRKKSNAK-----LATGAHV--GNVLVHESAKIGEGCLIGPDVAIGPCV 280
Db 236 VNRML-SEAKALQQRINRVMENGVTIIDPSFTFGTDVKIGIDTTIEFGVRIAGHTT 295

Qy 281 VEDGV-----RLSRCTVMR3VRICKHACISNSIIGWHSTYGVOWARIENMTILGEDVHVC 334
Db 296 IEEVDWIGQSYSEINNSTHISNANIKQ-SVINDSIVGENTTVGPFAQLRPGSNLGSVKV- 353

Qy 335 DEVSYNGGVVLPHKEIKSILK 356
Db 354 -----GNFV-----EVKKADIK 365

RESULT 11
US-09-370-838-190
; Sequence 190, Application US/09370838
; Patent No. 6444425
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Roach
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS OF
; FILE REFERENCE: 210121.475C1
; CURRENT APPLICATION NUMBER: US/39/370,838
; CURRENT FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: US 39/285,323
; EARLIER FILING DATE: 1999-04-02
; NUMBER OF SEQ ID NOS: 289
```

```
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 190
; LENGTH: 146
; TYPE: PRT
; ORGANISM: Homo sapien.
US-09-370-838-190

Query Match      12.8%; Score 239; DB 4; Length 146;
Best Local Similarity 35.9%; Pred. No. 1.3e-18;
Matches 47; Conservative 23; Mismatches 33; Indels 28; Gaps 2;

Qy 251 VVGNVLVHESAKIGEGCLIGPDVAIGPCVVEDGVRLSRCTVMRGVRIKKGACISNSIIG 310
Db 9 LLGNVYIHTAKVAPSAVLGNPNVSGKVTGEGVRLRESIVLHGATLQEHCTVLHSIVG 68

Qy 311 WHSTVGOWARIEN-----MTILGEDVHVCDEVYSNGG 342
Db 69 WGSTVGWRWARVEGTSDDPNPDPRAPYDSESLFKGKLLPAITILGCRVRIPAEVLIILNS 128

Qy 343 VVLPHEIKSS 353
Db 129 IVLPHEKLSRS 139

RESULT 12
US-09-199-637A-175
; Sequence 175, Application US/09199637A
; Patent No. 6355411
; GENERAL INFORMATION:
; APPLICANT: Ausubel, Frederick
; APPLICANT: Goodman, Howard M.
; APPLICANT: Rahme, Laurence G.
; APPLICANT: Mahajan-Miklos, Shalira
; APPLICANT: Tan, Man-Wah
; APPLICANT: Cao, Hui
; APPLICANT: Drenkard, Eliana
; APPLICANT: Tsongalis, John
; TITLE OF INVENTION: VIRULENCE-ASSOCIATED NUCLEIC ACID
; FILE REFERENCE: 00786/365002
; CURRENT APPLICATION NUMBER: US/09/199,637A
; CURRENT FILING DATE: 1998-11-25
; PRIOR APPLICATION NUMBER: 60/066,517
; PRIOR FILING DATE: 1997-11-25
; NUMBER OF SEQ ID NOS: 437
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 175
; LENGTH: 224
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-199-637A-175

Query Match      12.3%; Score 228.5; DB 4; Length 224;
Best Local Similarity 33.3%; Pred. No. 3.7e-17;
Matches 75; Conservative 39; Mismatches 90; Indels 21; Gaps 8;

Qy 1 MKALILVGGFGRRLRLTLSPKPLVDFAKPMILHQIEALKEVGVTEVVLAINRPEVM 60
Db 1 MKAMILAGRGGERMRPTTLHTPKPLIEAAGVPLIERQLLALRQAGVDDVINHAWLGEQI 60

Qy 61 INFLKDFEDKLGITITCQETEPGLTAGPLALARDKLAGSQGPFFV-NSDVISPEPAE 120
Db 61 EAYLGD-GSRLGGR-AYSPEGEPLTGGITFRALPLLGE---QPELLNGDVWSDFDYSR 116

Qy 121 LIKPHKCHGGEATIMVTKVDEPSKYGVVME-EATGRVERFVEKPKIFVGKINAGIYLL 179
Db 117 L---H-ADGDL-AHLVL--VDKPAHPAGCFHLDAAGRVGETREAG---GNLTVSGIAVL 167

Qy 180 NPSVLDRIELRPTSTEKEVFOIAADQ---Q-LYAMVLPGFMDVVG 221
Db 168 HPALFEGCGPGAFKAPLLRLKRAIAAGRVSGEHYR---GQWVDVG 208
```

```
RESULT 13
US-09-252-991A-17377
; Sequence 17377, Application US/C9252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: 107196.136
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US/09/252.991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17377
; LENGTH: 258
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17377

Query Match 12.2%; Score 227.5; DB 4; Length 258;
Best Local Similarity 33.3%; Pred. No. 6.1e-17;
Matches 75; Conservative 39; Mismatches 90; Indels 21; Gaps 8;

QY 1 MKAILVGGFTRLRPLTSLSPKPLVDPAKPKMILHQIEALKEVGVTEVVLAINYRPEVM 60
DB 35 MKAMTLAAGRGMRPTTLHPKPLIEAAGVPLIERQLLALRQAGVYDWINHAWLGEQI 94
QY 61 INFLDFEDKLGITITCSOETPLCTAGPLALARDKLADGSGQPPFVLNSDVISEYPPAE 122
DB 95 EAYLGD-CSRIGGRYAYSPSEPLETGGIFRALPLGE---QFFLLNGDVNSDFIYSR 150
QY 121 LKFKHKHGGGATIMVTVDSPSKYGVVYMB-EATGRVEREVEPKIFVGNKINAGIYLL 179
DB 151 L---HLADGDLAHLVL--VDNPAHPAGDFHLDAGRGVETREAG---GNLTYSGLAVL 201
QY 180 NPSVLDRLRLPTSTKEVFPQIADQ---QLYAMVLPGFMMVVG 221
DB 202 HPALFEGCGQAGFLAPLRLKAIAGRVSGEYH----GQWVDVG 242

RESULT 14
US-09-107-532A-6010
; Sequence 6010, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
```

```
NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 6010:
SEQUENCE CHARACTERISTICS:
LENGTH: 393 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYDROTHERICAL YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (8) LOCATION 1...393
SEQUENCE DESCRIPTION: SEQ ID NO: 6010:
US-09-107-532A-6010

Query Match 11.5%; Score 214; DB 4; Length 393;
Best Local Similarity 25.0%; Pred. No. 3.8e-15;
Matches 92; Conservative 60; Mismatches 150; Indels 66; Gaps 15;

QY 3 ALLILVGGFTRLRPLTSLSPKPLVDPAKPKMILHQIEALKEVGVTEVVLAINYRPEVMIN 62
DB 18 AIIILVGGFTRLRPLTSLSPKPLVDPAKPKMILHQIEALKEVGVTEVVLAINYRPEVMIN 70
QY 63 FLKDFEDKLGITITCSOETPLCTAGPLALARDKLADGSGQPPFVLNSDVISEYPPAE 120
DB 71 ---QVKAQLGERSKVALQAEQLGT--GHAVLQAASFLEGGKGTTLVISGDTPLLTETLNN 126
QY 121 LKFKHKHGGGATIMVTVDSPSKYGVVYMBEATGRVEREVEPKIFVGNKINAGI 176
DB 127 LFEYHGGKNASATILTAAGNPTGYGRIIRGH-IGIVEKIVEQKDATPEALVQEIINTGT 185
QY 177 YLL-NPSVLDRLRLPTSTKEVFPQIADQ---QLYAMVLPGFMMVVG 222
DB 186 YCFDNEALFDALSKVGTNNAGQYYLTDIILKEEGHTVAAYQ-----TDDFEESMG- 238
QY 223 PRDYITGLRLYLDSDIRKXSAALATGAHVGNVNLVHESAKIGEGCLIGPOVAIGPCVVE 282
DB 239 VNDRIALAK--ANEIMRKRIQM---HMVANGVSFVDSAT---TYIDAGVEIGPDTLIE 288
QY 283 DGVRLSRCTVMRG-----VRKKKACISNII-----GWHSTVQWARIENMT 325
DB 289 AGVQIQGNTVGSDCVIGSHSKIVDSRIEDHVVIENSVISSHVKKHADVGPYAHLRPKA 348
QY 326 ILGEDVHV 333
DB 349 EIGENVHI 356

RESULT 15
US-08-469-202-14
; Sequence 14, Application US/08469202
; Patent No. 5750875
; GENERAL INFORMATION:
; APPLICANT: STALKER, DAVID
; TITLE OF INVENTION: GLYCOCEN BIOSYNTHETIC ENZYMES IN PLANTS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CALGENE, INC.
; STREET: 1920 FIFTH STREET
; CITY: DAVIS
; STATE: CA
; COUNTRY: USA
; ZIP: 95616
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 2.0 MB
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 7.1
```

```
/ SOFTWARE: Microsoft Word 5.1 (a)
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/469,202
/ FILING DATE: 6-JUNE-95
/ CLASSIFICATION: 800
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/016,881
/ FILING DATE: 11 FEB 1993
/ CLASSIFICATION: 800
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Elizabeth Lassen
/ REGISTRATION NUMBER: 31,845
/ NAME: Donna E. Scherer
/ REGISTRATION NUMBER: 34,719
/ NAME: Carl J. Schwedler
/ REGISTRATION NUMBER: 36,924
/ REFERENCE/DOCKET NUMBER: CGNE 93-2
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 916-753-6313
/ TELEFAX: 916-753-1510
/ INFORMATION FOR SEQ ID NO: 14:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 431 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-469-202-14

Query Match      10.4%; Score 193; DB 1; Length 431;
Best Local Similarity 22.2%; Pred No. 9.9e-13;
Matches 94; Conservative 61; Mismatches 170; Indels 98; Gaps 15;

QY 3 ALILVGGFGRTRLPRLTSLFPPKPLVDFAFKPMIL-HQIEALKEVGVTGVVLAINYRPEVMI 61
DB 1111111111111111111111111111111111111111111111111111111111111111
DB 22 ALILAGGRGTRLKDLTKRAKPAVHFGGKFRIDFALSNCINSGIRRMGVITQYQSHTLV 81
QY 62 NFLK-----DPEDKLGIITCSQETEPGLTAGPLALARDKLAGDGGQPPFFVLNS 110
DB 82 QHIQKQMSFFNEEMNEFVDLPAQORMKGENRYGTDAVTONLQIIRRYKAEYVVLG 141
QY 111 DVISEYPPFAELIKFKHCHGGEATI--MVTKVDEPSKYGVVWMBEATGRVERFEKPKIFV 168
DB 142 DHIIYQDYSRMLIDHVEKGARCTVACMPVPIEASAFGVMAVDE-NDKIIEFVEKP--- 196
QY 169 GNK-----INAGIYLLNPSVL-----DRIELRPTSTIEKEVFPQIAADQQLYA 210
DB 197 ANPPSPNDPSKSLASMGIVYFDADLYELLEDDDDSDNSHDFGKOLIPKITEAGLAYA 256
QY 211 MVLP-----GFWMGVQPRDYITG-----LRLYLDLSIRKK 240
DB 257 HPFFPLSCVQSDPDABFYWRDVGTLKAYWKANLDLASVVFELDMYDRNWPPIRTYNESLPPA 316
QY 241 SAAKLATGAHVVGNYLVHESAKIGEGCLIGPDVAIGPCGVVEDGVRLSRCTWMRGVRIKK 300
DB 317 KPVQDRSGSH--GMTL--NSILVSDGCVIS-----GSVVVQSVLFSR-----VRVNS 358
QY 301 HACISNSIIGHWSTVGQWARIENMTI-----LGEDVHVCDVY--SNGGVWLPH 347
DB 359 FCDIDSAVLLPEVWVGRSCRRCRVIDRACVPIEGMWIGENNAEDARRFYRSEEGIVLVT 418
QY 348 KEI 350
DB 419 REM 421
```

Search completed: October 9, 2003, 14:04:07  
Job time : 22 secs